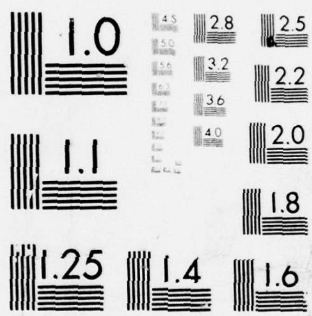
D-A061 646 NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA F/G 20/6
GREY-SCALE VERSUS COLOR CODING OF ACOUSTIC DATA IMAGES.(U)
MAR 78 R S FRENCH
UNCLASSIFIED NOSC-TR-207 NL

1 OF 1
AD
A061646

END
DATE
FILMED
2--79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL

(12)

AD A061 646

NOSC

NOSC TR 207

Technical Report 207

GREY-SCALE VERSUS COLOR CODING OF ACOUSTIC DATA IMAGES

RS French

31 March 1978

Final: October 1976 to September 1977

Prepared for
Naval Electronic Systems Command

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

ADMINISTRATIVE INFORMATION

The display and operator procedures project is sponsored by the Naval Electronic Systems Command, ELEX-320, under the block program in undersea surveillance, program element 62711N, subproject XF-11-101-100. The work reported here was performed during the period October 1976 through September 1977.

The author gratefully acknowledges the critical role of Ronald Domb and John Hammond who were responsible for the development of the software and data sets for these experiments, the technical support and review provided by Dr. Charles Persons and Dean Hanna, and the important contribution of seven members of the Display Branch who served enthusiastically as subjects in these experiments.

Released by
DA Hanna, Head
Signal Processing and
Display Division

Under Authority of
HA Schenck, Head
Undersea Surveillance
Department

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER: NOSC TR-207
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 GREY-SCALE VERSUS COLOR CODING OF ACOUSTIC DATA IMAGES,
5. TYPE OF REPORT & PERIOD COVERED: 9 Final rept. October 1976 - September 1977,
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 Robert S. French
8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Ocean Systems Center, San Diego, CA 92152
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62711N, XF 11-101-100
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Electronic Systems Command, Washington, DC 20360
12. REPORT DATE: 11 31 March 1978
13. NUMBER OF PAGES: 83
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 16 F11101 17 XF11101100
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited. 12 64p.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): pattern recognition, images, acoustics, colors, coding

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Acoustic data images were studied using both grey-scale and color coding for 8 to 64 discriminable levels of information. Display image encoding using 8 discriminable levels of grey-scale or color was shown to provide good perception of pattern features. Beyond 8 levels, it was increasingly more difficult to distinguish individual levels and pattern perception was degraded. With 32 to 64 levels of quantization available, excellent pattern perception was achieved using a color "band-step" coding scheme. The color bands code the major levels of amplitude, while the intensity gradient within each band highlights contouring and provides for finer amplitude discrimination. The use of six color bands of 5 to 10 intensity steps within

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

393 159 78 11 24 070

20. Continued.

each band was particularly effective. Color displays were shown to be superior for the perception of patterning, contouring, and equivalent amplitudes and also to be equivalent to grey-scale displays for discriminating patterns of very weak signals in noise.

SUMMARY

PROBLEM

Color displays have proven to be an effective medium for studying the characteristics of a class of acoustic data images characterized by strong patterning. However, black-and-white displays have the advantages of greater availability and reliability, lower cost, and inherently higher resolution. Consequently, a quantitative comparison of the relative perceptibility of signal-related patterning between black-and-white and color displays seemed necessary.

RESULTS

Color displays were shown to be superior for the perception of patterning, contouring, and equivalent amplitudes and also to be equivalent to grey-scale displays for discriminating patterns of very weak signals in noise.

Display image encoding using 8 discriminable levels of grey-scale or color was shown to provide good perception of pattern features. Beyond 8 levels it was increasingly more difficult to distinguish the individual levels and pattern perception was degraded. With 32 to 64 levels of quantization available, excellent pattern perception was achieved using a "band-step" coding scheme. Band-step coding, for example, may utilize 6 color bands with 5 to 10 intensity steps within each color band. The color bands code the major levels of amplitude, while the intensity gradient within each band highlights contouring and provides finer amplitude discrimination.

CONTENTS

CONTENTS (Continued).

ILLUSTRATIONS

ILLUSTRATIONS (Continued).

TABLE

INTRODUCTION

BACKGROUND

The Comtal color display (see appendix A) has proven to be an effective medium for reviewing and photographing a class of acoustic data images that is characterized by strong patterning. Examples of images of this type may be found throughout this report. The focus of the series of experiments reported here is a comparative evaluation of various ways of coding the amplitude values of strongly patterned data images in grey-scale or color. This area of concern differs from many of the previous studies comparing color and grey-scale coding of acoustic data.[1,2] The task investigated here is the perception of patterning which is considerably above the visual threshold and thus readily apparent. The studies cited, on the other hand, are concerned primarily with the threshold of detection of signal energy within a noise field. The present experiments are somewhat more analogous to the studies reported by Christ[3] dealing with easily resolvable stimuli such as letters in a noiseless field and by Lamar, et al.,[4] dealing with grey-scale/color comparisons of Skylab images.

Various coding schemes have been developed and evaluated informally over a period of several years in the NOSC Surveillance Display Laboratory. A coding scheme consisting of six color bands each containing an intensity ramp of 10 steps has proven in practice to be a more effective way of coding amplitude for this class of images than has any alternative which has been tried. Grey-scale equivalents for such coding have not proven to be as good because the higher intensities are difficult to resolve. Display of such images may be an issue in the design of future systems. There is a need, therefore, for quantitative information on the effect of coding on the perceptibility of pattern features in such images. In particular, there is a need for information on the feasibility of substituting grey-scale coding for color; moreover, information is needed on which type of coding scheme is best. The issue of the grey-scale versus color coding is of practical interest in view of the fact that black-and-white display consoles are readily available, reliable, inexpensive, and have inherently higher resolution than color consoles. A related issue investigated in these experiments is the value of interactive control of the threshold and gain quantization parameters by the operator. Quantitative information on any benefit from interactive coding can help determine the need for this option in future system procurements.

---

[1] Butler, WM, "Color and Black and White Display of Sonar Information," Tracor Document T71-AU-9621-U, 25 February 1971.

[2] Butler, WB, and McKemie, WM, "Engineering Guidelines for the Use of Color in a Sonar Display," Tracor Document T73-AU-9550-U, 1 April 1974.

[3] Christ, RE, and Corso, GM, "Color Research for Visual Displays," ONR Report No. ONR-CR213-102-3, July 1975.

[4] Lamar, DL, and Merifield PM, et al., "Evaluation of Pseudocolor Transformations of Skylab and Landsat Images and Test Charts, Part V, Final Report; SKYLAB EREP463 Results," CalESCO Technical Report 76-1, 1976.

OBJECTIVE

The overall objective of the series of four experiments reported here was to determine quantitatively the relative perceptibility of signal-related patterning using various grey-scale and color schemes that are effective for their intended use and better than existing alternatives.

APPROACH

GENERAL

In broad terms the approach taken to compare coding schemes was to obtain judgments from a team of "experts" on the relative perceptibility of acoustic data pattern features when the images were viewed on the Comtal display system using selected grey-scale and color coding (or quantization) schemes. Measures of relative perceptibility were obtained from four related experiments.

In Experiment 1, relative ranking was obtained for six candidate coding schemes ranging from 8 to 64 levels of amplitude quantization. The rankings were made independently against four perceptibility criteria. Although both grey-scale and color coding was included, no attempt was made to obtain a direct comparison of perceptibility in this experiment. Experiment 1 provided a preliminary indication of coding preferences and eliminated from further consideration any coding schemes that were clearly unsuitable.

In Experiment 2, data were obtained on the preferred settings of threshold and gain (adjustment of quantization cutoff and range relative to the range of the data). One of the top-ranking coding schemes from Experiment 1 was used in this experiment.

In Experiment 3, the relative perceptibility of equivalent grey-scale and color coding schemes were compared in terms of the threshold of pattern perceptibility measured in decibels. A calibrated test set of acoustic data patterns was used in which the signal-to-noise ratio (SNR) was controlled by adding Gaussian noise incrementally. Comparisons were made using 8-, 16-, and 60-level schemes which had been compared initially in Experiment 1. The preferred threshold and gain values obtained from Experiment 2 were used to requantize the calibrated test data set.

Finally, in Experiment 4 the subjects were asked to make direct ranking judgments of the relative perceptibility of grey-scale versus color-coded data patterns when quantized in 8-, 16-, and 60-levels and against the same four criteria used in Experiment 1.

CODING SCHEMES

The six coding schemes that were compared in Experiment 1 are conveniently described as combinations of numbers of bands and intensity steps within each band. A brief discussion on this concept as it applies to the quantization and coding of data for the Comtal display is provided in this section.

The Comtal display has the capability for 6-bit (64 level) data coding in either grey-scale or color. Coding schemes were included in Experiment 1 which would be appropriate to display systems limited to either 3 bits (8 levels) or 4 bits (16 levels) of data coding. The six coding schemes selected for comparison are illustrated by the photographs in figure 1.

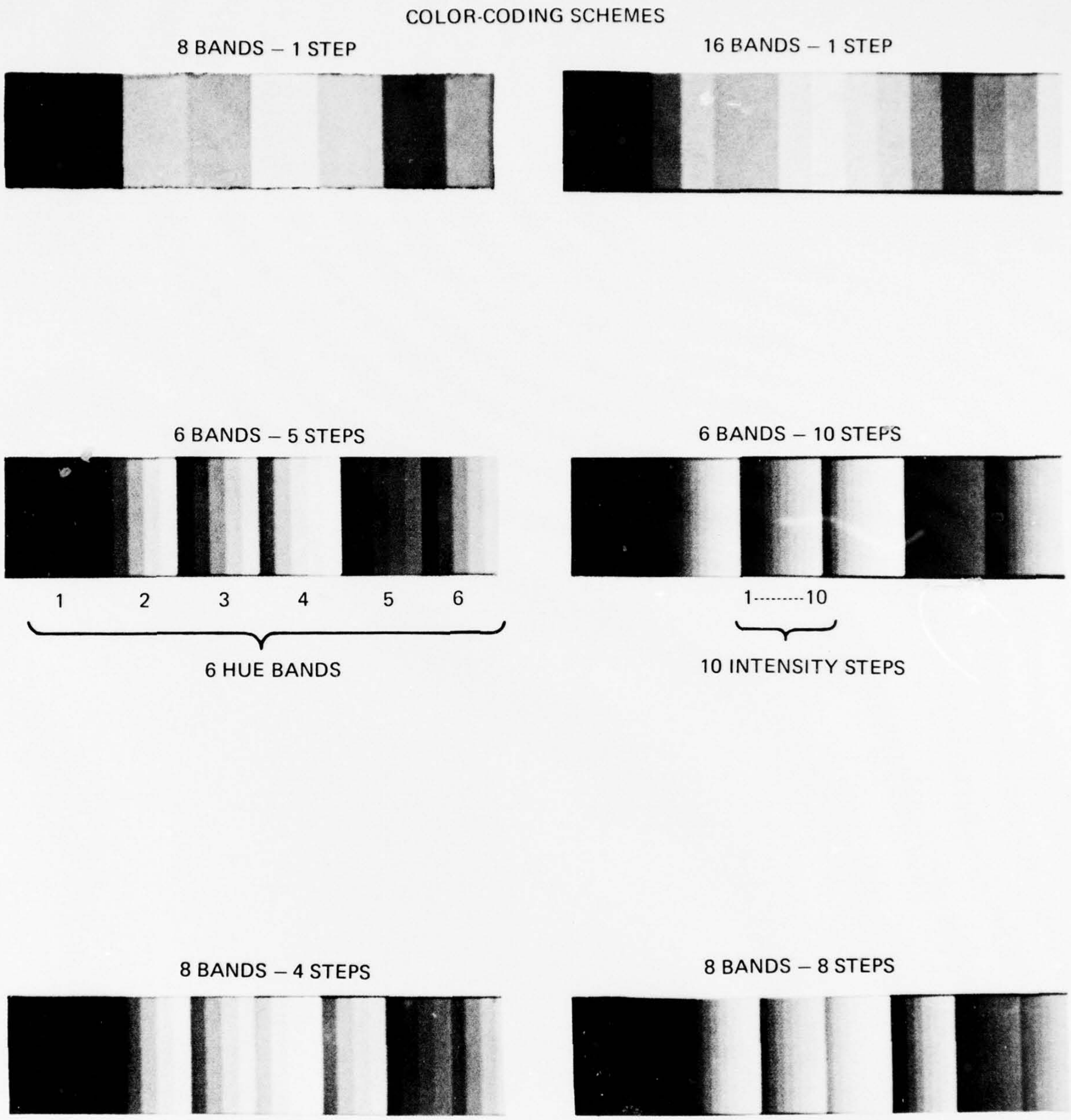
All data images included in this report are based on an array which consists of 256 data points in the horizontal and 80 data points in the vertical. The vertical array has been replicated by a factor of three and smoothed by simple linear interpolation to produce a display of 256 X 240 points.

Note from figure 1 that the coding schemes selected consist of 6, 8, or 16 bands in combination with 1 to 10 steps within the bands. In general the coding schemes using hue are based on a series of bands each in a different color (hue). The choice of hue does not appear to be critical, except that they should be easily distinguished between adjacent bands and represent a logical progression from cool colors (e.g., blue) to hot colors (e.g., red and pink). With the exception of the one-step schemes, each band is composed of a series of graduated intensity (or brightness) steps. The effect of this band-step coding is to produce concentric contours in the images as may be seen in figure 2. This contouring effect is produced because there is an abrupt intensity change between the highest level of a given hue and the lowest level of the immediately adjacent hue. Band-step coding effectively codes major differences in amplitude in different easily recognized hues; while fine-grain differences in amplitude are revealed as contours of the same hue which differ slightly in intensity.

Some interesting features of band-step coding in color compared to grey-scale are brought out by figure 3. A grey-scale being unidimensional can vary in brightness only; a color scale, however, can vary in both hue and brightness. Accordingly, in the construction of the color-coded schemes, as illustrated in figure 3(b), the 8-band, 1-step scheme uses brightness and hue to achieve both a finer grain amplitude discrimination and the contouring effect. With grey-scale coding as illustrated in figure 3(a), the only variable available for achieving discriminable levels is brightness or grey-scale. Because the response of the human eye to brightness differences is not linear it is difficult to construct a grey-scale that achieves the same discriminability of the higher amplitude bands as does the comparable color scheme.

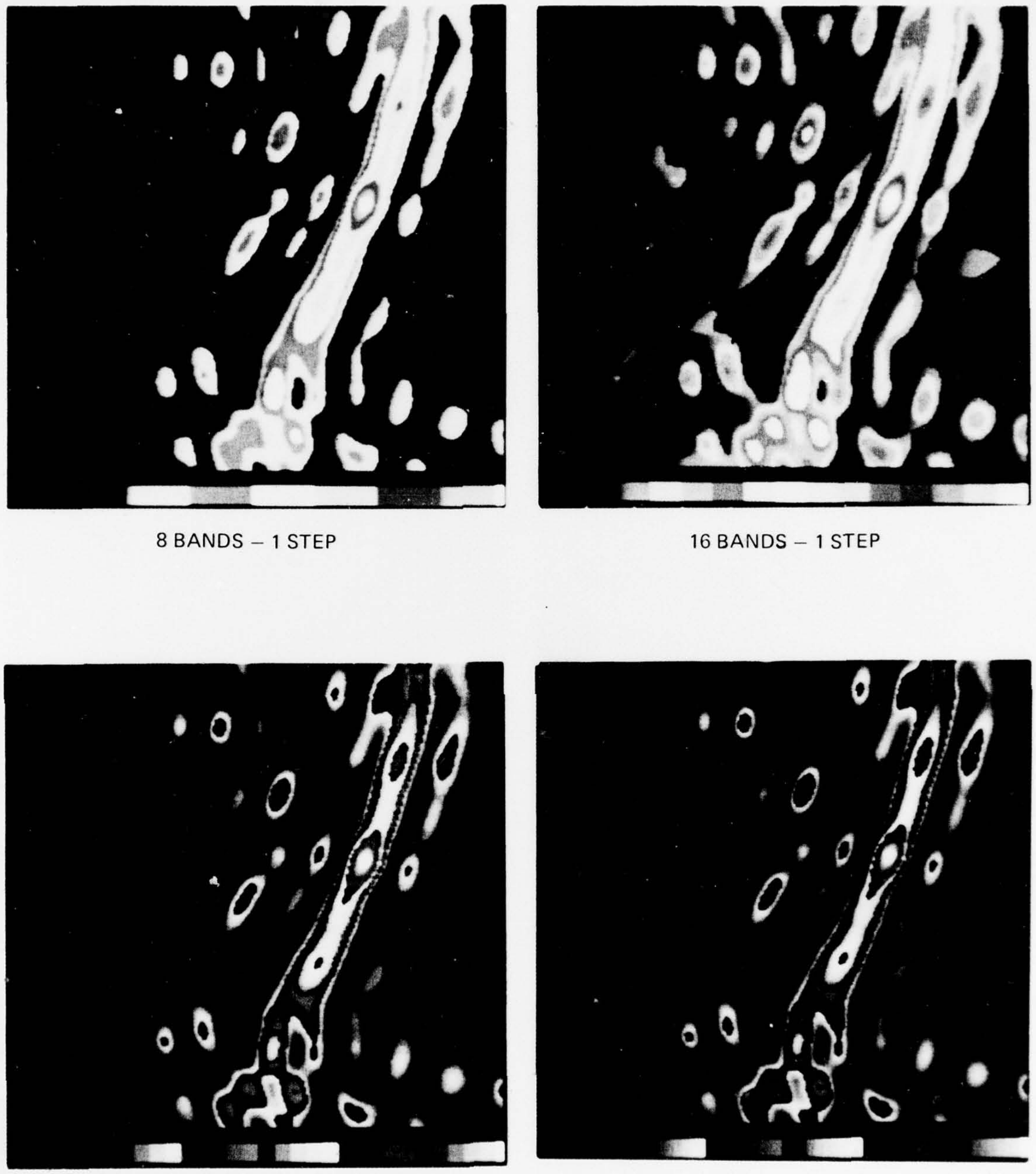
The effect of different levels of quantization in grey scale coding may be seeen from the photographs in figure 4. Note that as the number of levels is increased from 8 to 16 the image appears to be out of focus. The reason for this is that with the increase in coding levels the difference in intensity between adjacent levels is smaller and is not easily resolved by the eye at the higher intensities. Although 64-level coding in either grey-scale or color can be implemented on the Comtal, the resultant images appear even more out of focus than the 16-level images (at least for the data used in these experiments).

To achieve nearly equal increases in brightness requires a logarithmic increase in intensity or luminance. To make the higher intensity bands discriminable, it was necessary to increase the size of the steps relative to the lower intensities, particularly when the total number of levels was from 32 to 64. In general, the resulting quantization functions consisted of two or three linear segments with progressively greater slope. The functions finally chosen were developed by the "cut and try" method. The objective of this process was to retain as much discrimination as possible at the critical higher amplitudes while retaining some perception of the lowest levels. The task was complicated by the requirement to maintain the same number of intensity steps within each band.



NOTE: THE BANDS CONSIST OF EASILY DISTINGUISHED HUES. AN INTENSITY GRADIENT OF A SPECIFIED NUMBER OF STEPS (FROM 1 TO 10) IS CONSTRUCTED WITHIN EACH BAND.

Figure 1. The six coding schemes compared in Experiment 1 with color coding.



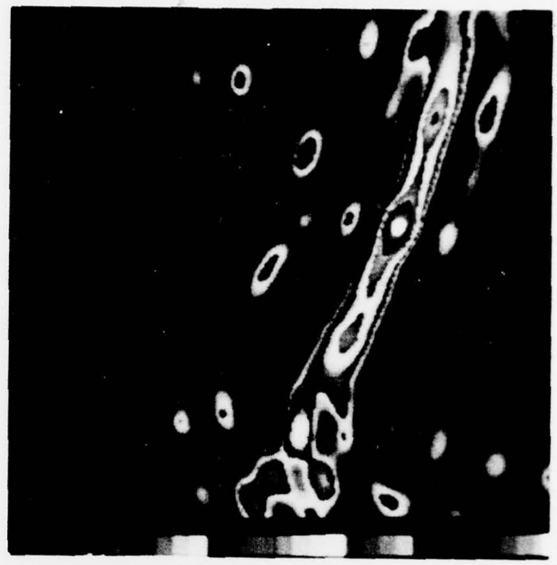
8 BANDS – 1 STEP

16 BANDS – 1 STEP

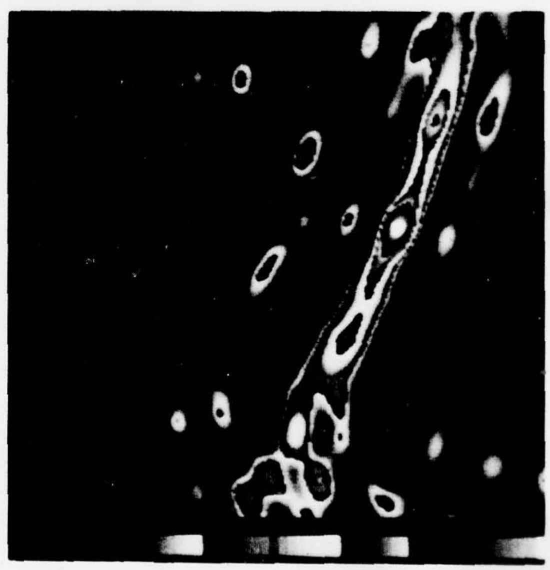
6 BANDS – 5 STEPS

6 BANDS – 10 STEPS

Figure 2. The color coding of the six schemes.



8 BANDS – 4 STEPS



8 BANDS – 8 STEPS

Figure 2. (Continued).

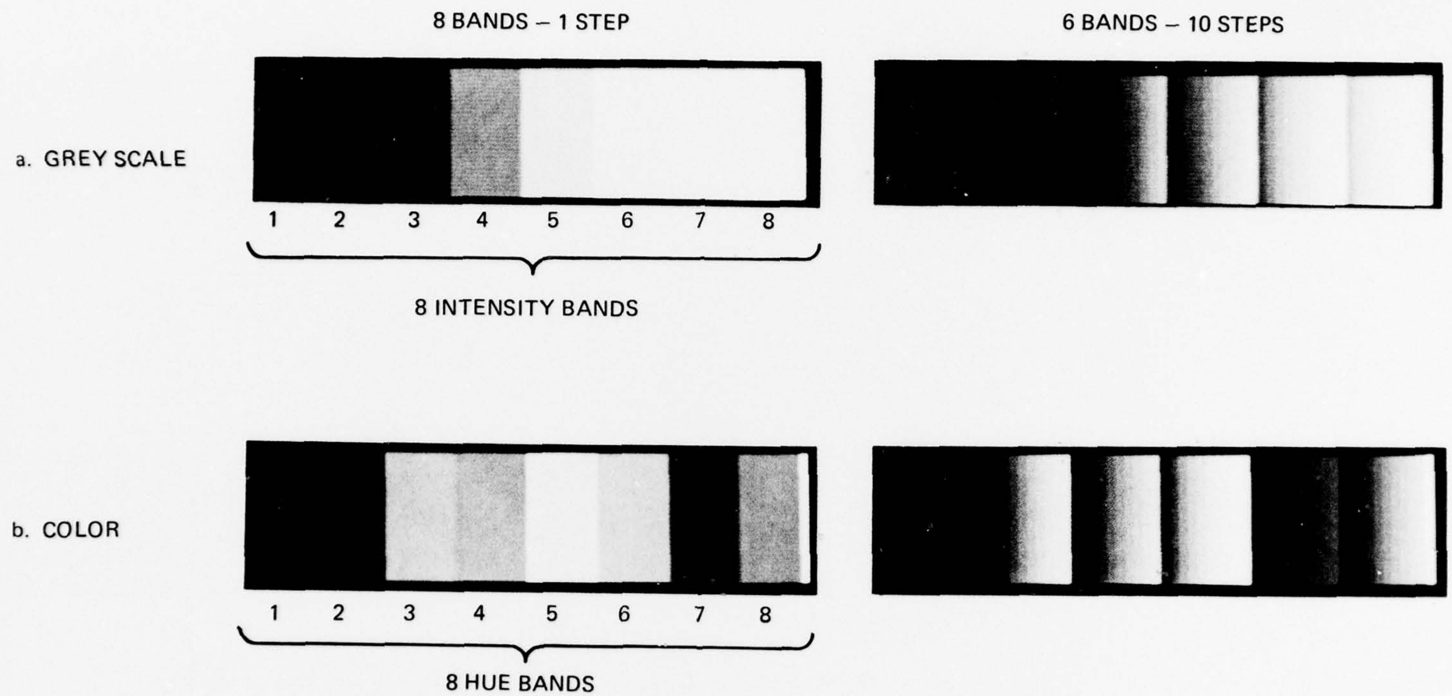
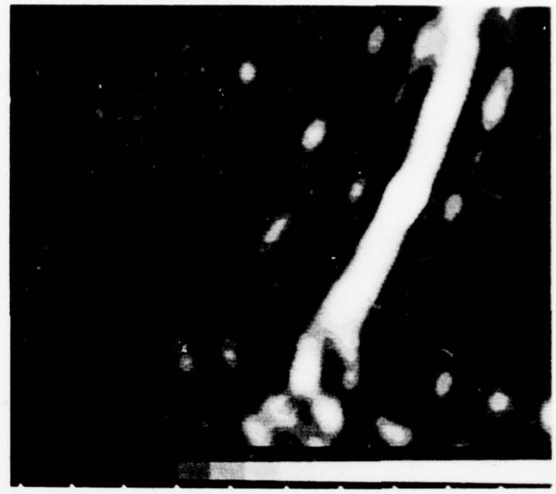
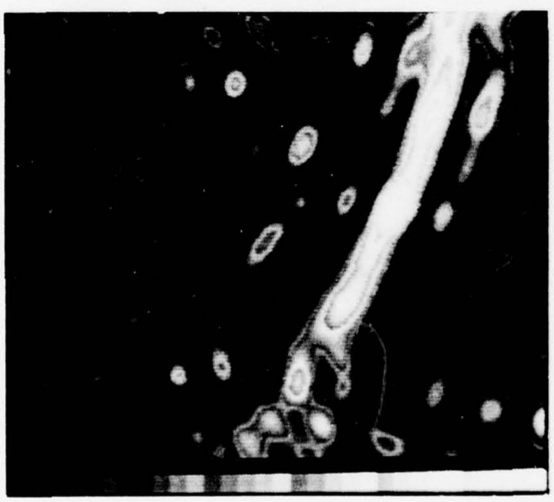


Figure 3. Features of band-step coding. With only one step (intensity level) within a band, solid bands of color are used to approximate grey levels.

8 BANDS – 1 STEP



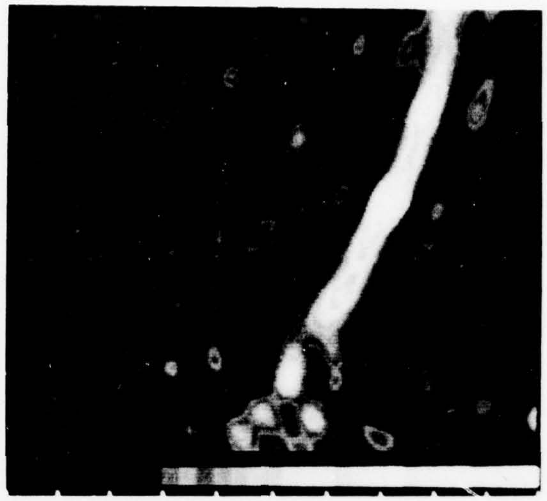
16 BANDS – 1 STEP



6 BANDS – 5 STEPS

6 BANDS – 10 STEPS

Figure 4. The six grey-scale coding schemes.



8 BANDS – 4 STEPS

8 BANDS – 8 STEPS

Figure 4. (Continued).

The grey-scale quantization ramps finally selected by this method are shown in figures 5 (a through f). In this set of figures, display brightness in arbitrary units from 0 to 64 is plotted as a function of data amplitude from 0 to 255 for the six coding schemes included in Experiment 1. The difference between the single- and multiple-step schemes is clearly evident from these figures as is the increase in the slope of the functions with increasing amplitude.

Luminance measurements were made for each of the 64 levels of Comtal grey-scale display brightness using the Model 2400 (Gamma Scientific, Inc.) Digital Display Photometer with a photo-optic filter. The resultant luminance function is shown in figure 6 with footlamberts on the ordinate and the 64 levels on the abcissa. Note that this function is approximately linear over a significant portion of its range. The shape of this function indicates that equal brightness differences can be achieved using the Comtal grey-scale, if values are selected as described above such that K log L_1/L_2 is approximately constant, where L_1 is a "just noticeable" difference and L_2 is the previous baseline.

CALIBRATED TEST DATA SETS

The development of calibrated data sets was an essential first step before it was possible to obtain data comparing perception thresholds with various coding schemes. This section discusses briefly the development and use made of these data sets.

A calibrated test data set such as shown in figure 7 is produced by adding calibrated amounts of Gaussian noise at an appropriate stage in the signal processing. Each data set consists of 11 images including an original image (without noise added) and an image consisting of pure noise. The original image is derived from a segment of data with a high signal-to-noise ratio (SNR). For this case, the SNR can be determined from the maximum amplitude of the image. Then, using this case as a reference, known amounts of noise power are added to the original signal plus noise to produce a readily determined change in SNR. The same additive noise waveform is used for each SNR, which is defined for the data in these experiments as 10 $\log_{10}$ signal power/total noise power in a 1-Hz frequency band at the input to the signal processing algorithm. The difference from one image to the next is the change in SNR which has been calculated in decibel units ranging from +5.8 dB to -30.0 dB in the case of the data in figures 2 through 7.

The signal-processing algorithm normalizes the signal-plus-noise power at the processor input to 1.0; this restricts the processor output amplitudes to the range of 0.0 to 1.0. If the SNR is very large the maximum output amplitude is near 1.0. Because the total power is held constant at 1.0, the maximum output amplitude decreases as the SNR decreases. At some point, which is dependent upon the processing bandwidth and integration time, the input noise predominates over the signal. Beyond this point, the maximum output amplitude is determined by noise alone, which remains constant as a result of the normalization. It may be noted from figure 7 that for the processing parameters in these test data the noise becomes predominant at approximately –6.0 dB. Beyond this SNR level the maximum output amplitude remains constant at approximately 0.5 (yellow coding).

Based on data sets such as these, measurement of the threshold of pattern perceptibility can be made for comparisons of various coding schemes. However, the concept of a threshold for the perception of distinctive pattern features requires some elaboration. It should be emphasized that a pattern threshold is distinctly different from the threshold for the

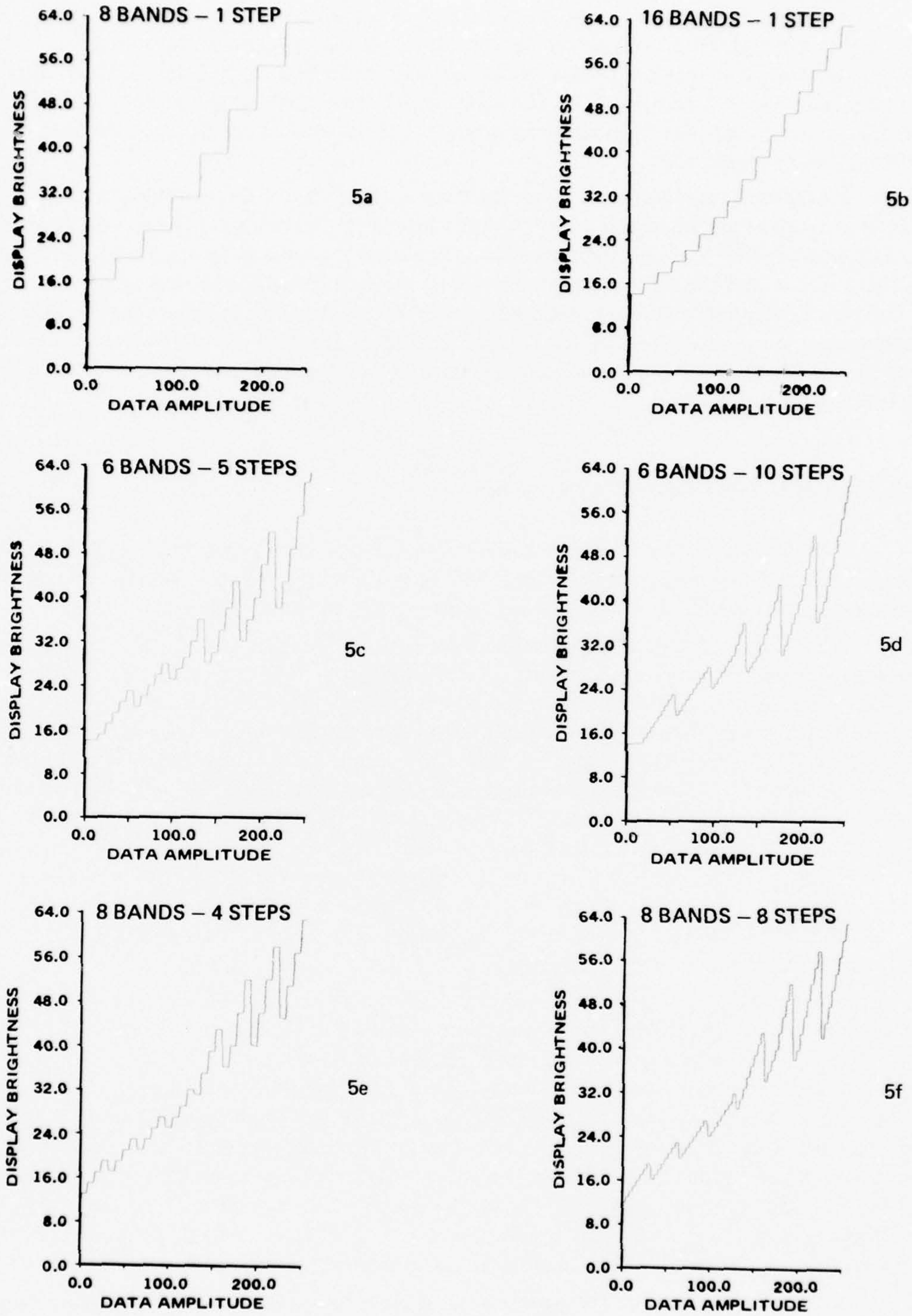


Figure 5. Grey-scale quantization ramps in the Comtal for the six coding schemes.

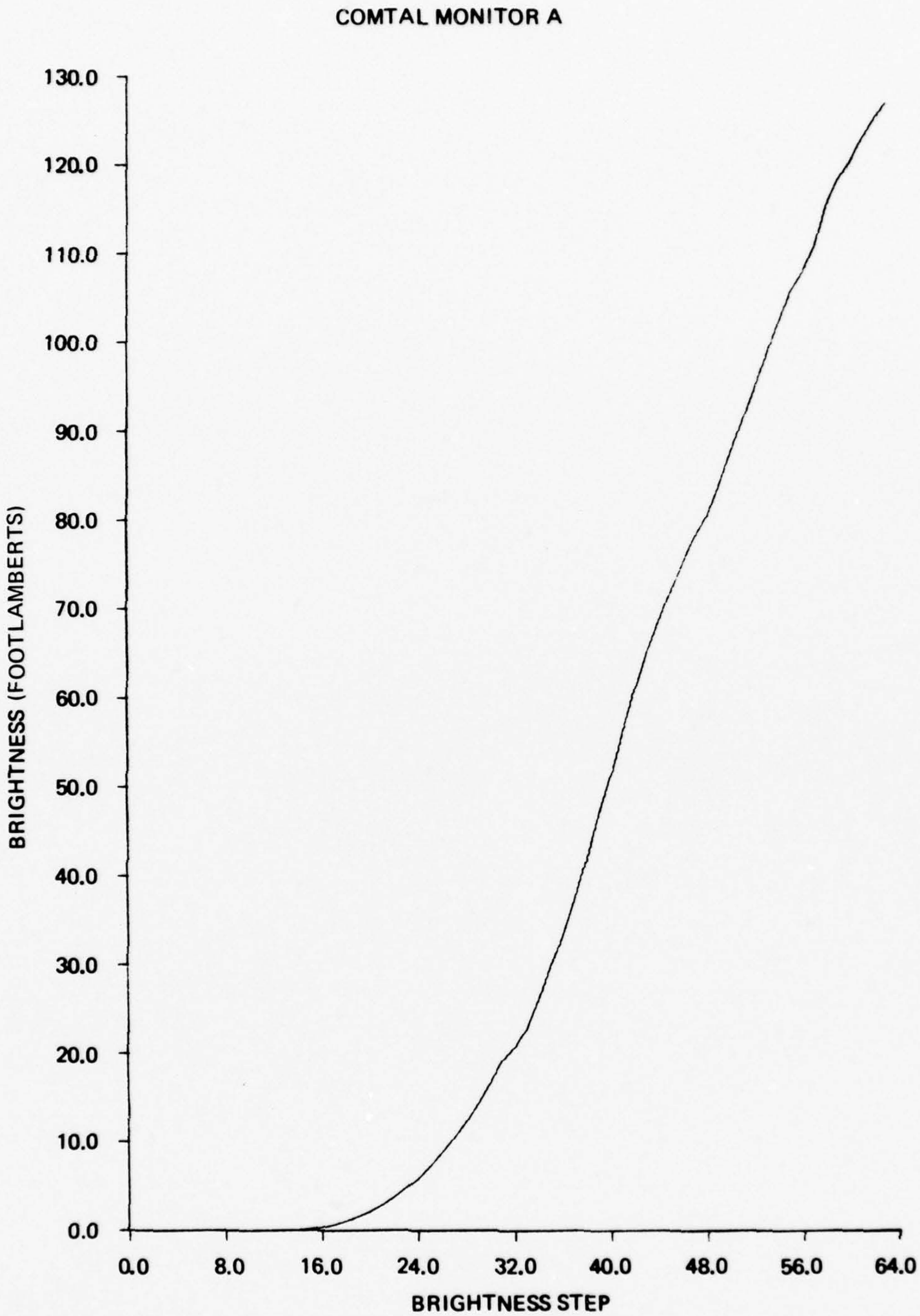


Figure 6. Luminance of the 64 levels of grey in the Comtal display system measured with a digital display photometer.

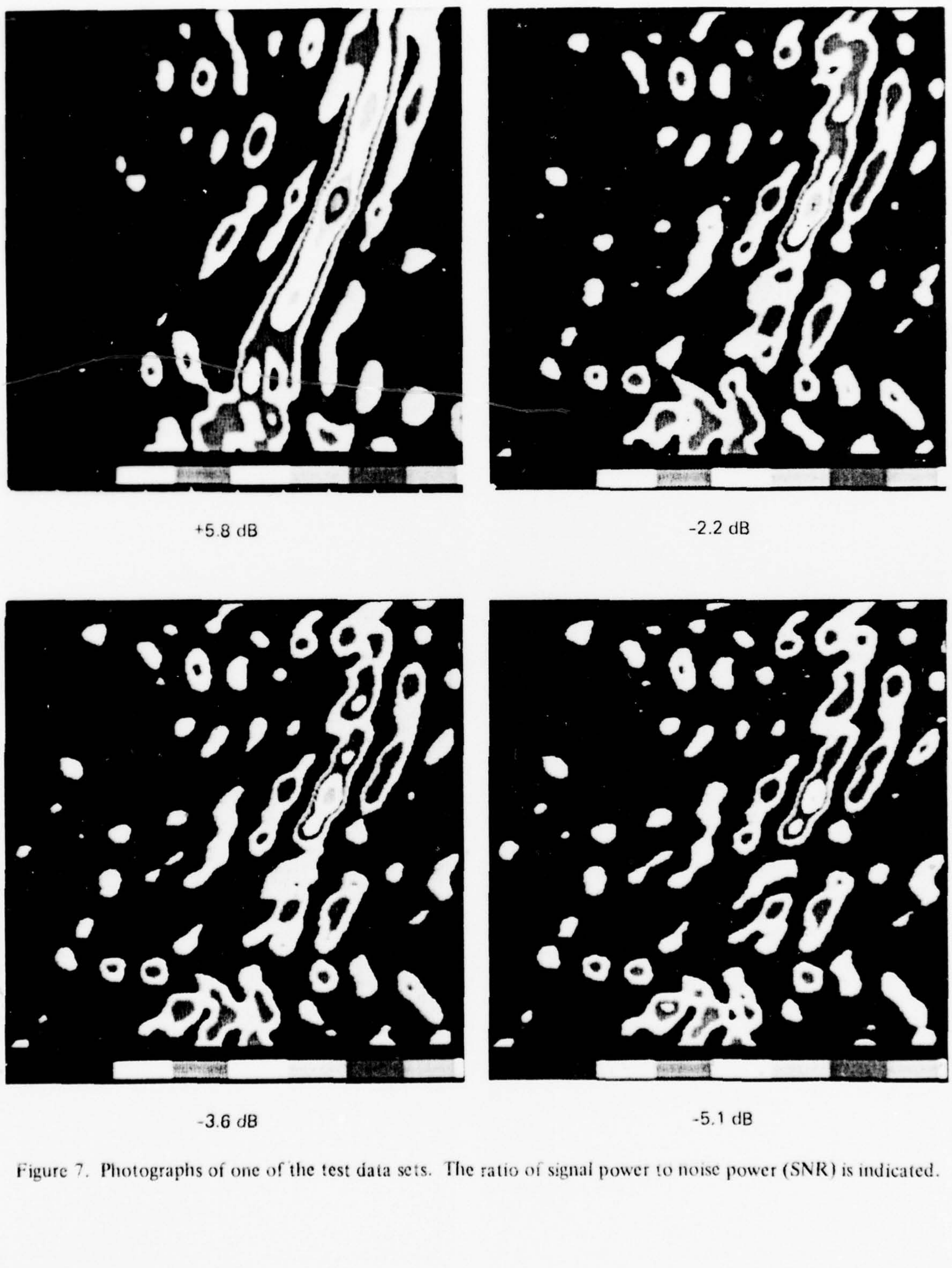


Figure 7. Photographs of one of the test data sets. The ratio of signal power to noise power (SNR) is indicated.

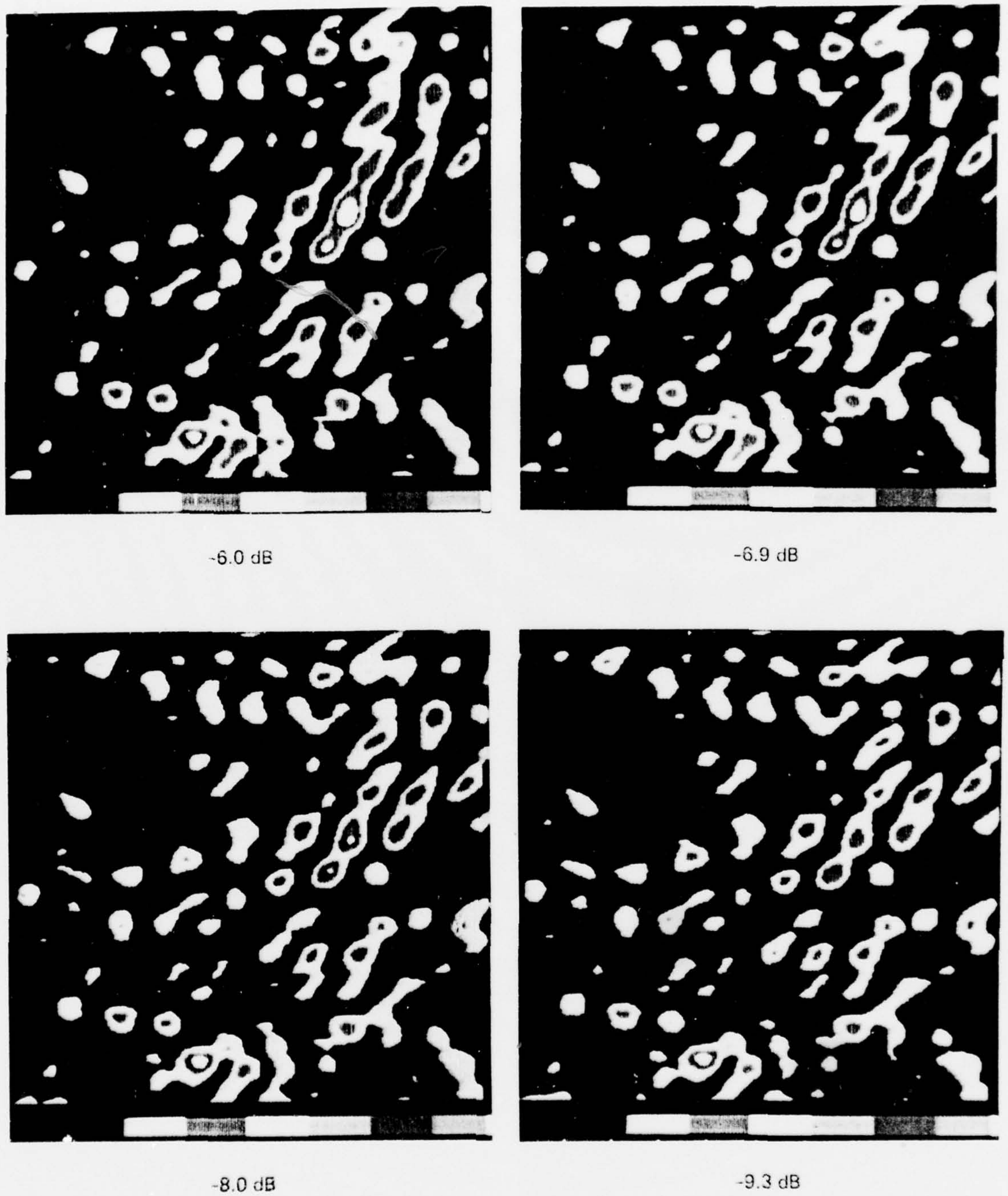
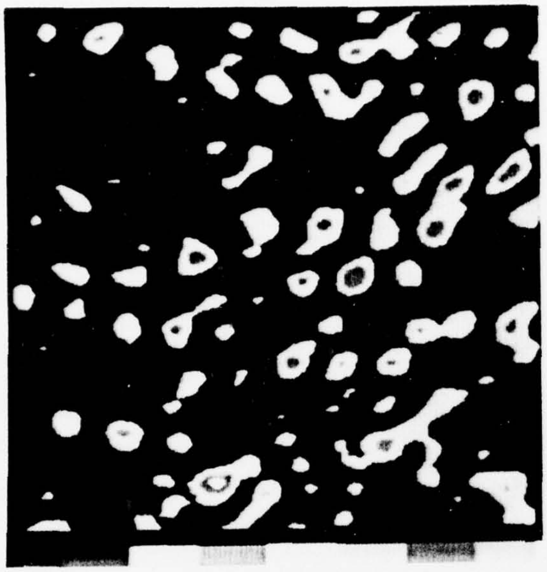
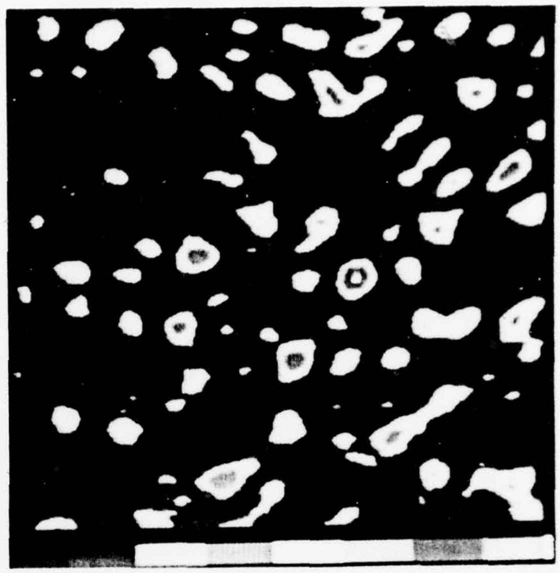


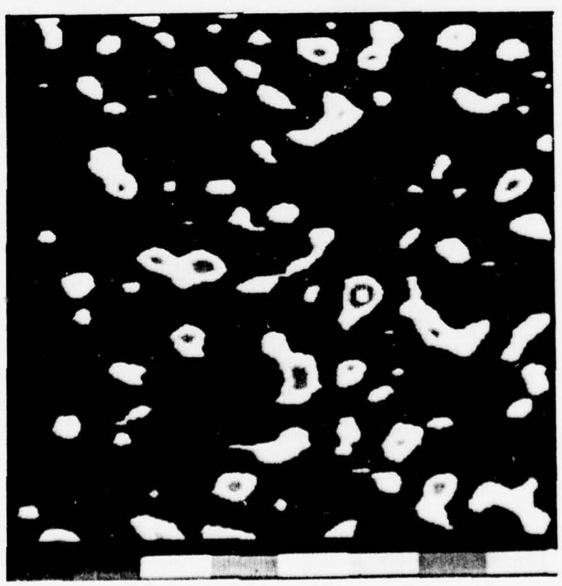
Figure 7. (Continued).



-11.4 dB



-15.5 dB



ALL NOISE

Figure 7. (Continued).

detection of a signal in noise. Clearly, the criteria for perception of a pattern are more subjective and less precise than for a tone or simple visual stimulus. The judgment of a pattern threshold is highly dependent on the viewer's understanding of the most significant pattern features and patterning in the noise distribution which might be confused with signal. It is expected that the threshold would vary as a function of the pattern features emphasized to the subject, the particular pattern viewed, and the subject making the judgment. The amount of practice in making similar judgments on other patterns and other coding schemes would also affect the measured threshold. It should be evident, therefore, that the concept of a pattern threshold and the comparison of thresholds under different conditions are statistical concepts. The approach is based on the assumption that sufficient consistency can be found in the threshold judgments of a group of subjects familiar with the type of patterns viewed to establish real differences, if any, between the thresholds for the various coding schemes.

The procedure used in Experiment 3 differs from conventional psychophysical measurement of a threshold value, primarily in the emphasis on full disclosure of all relevant parameters to the subjects. The subjects were allowed to examine in detail the entire range of SNR values in the two patterns and under all of the grey-scale and color-coding schemes prior to making any of their threshold judgments. The details on the procedures used to provide this exposure to the entire data set are included in Experiment 3. At this point, however, it may be noted that the procedure allowed the subjects to view images in the test series, four at a time on the Comtal and to page forward or backward through the entire series of 11 images as desired. This technique allows the subject to examine and compare pattern features in the original image and the noise image. With full knowledge of these two end points, he is better able to judge the SNR level at which the significant pattern features in the original image have been degraded beyond recognition. He is also better able to recognize patterns in the noise, which might otherwise be confused with the signal. The technique is based on the assumption that the subject will examine the test series carefully, several times in ascending and descending order, as he refines his judgment of the threshold. In a sense, his final judgment is equivalent to the mean of a series of independent threshold determinations obtained by presenting one image at a time in ascending and descending order.

THRESHOLD AND GAIN CODING ADJUSTMENT

The primary objective of Experiment 2 was to obtain preliminary information on the preferred manual settings of the threshold and gain coding parameters which may be adjusted by the operator in the Comtal display using the trackball. Briefly stated, the threshold and gain parameters determine the relationship between the dynamic range of the input amplitude scale and the output amplitude coding scale. This relationship is determined within the Comtal function memory by a transfer or mapping function. The potential value of having a threshold and gain adjustment is that the full dynamic range of the display output can be applied to any portion of the input amplitude distributions, where a finer discrimination of amplitude levels may be desired. As noted, the threshold and gain settings preferred by the subjects in Experiment 2 were desired as a basis for requantizing the test data sets for Experiment 3. A secondary objective of Experiment 2 was to compare the preferred manual settings of threshold and gain with those obtained by autonormalization. A brief discussion is provided of the threshold and gain coding parameters as they are implemented on the Comtal display.

The acoustic data used in these experiments consisted of a 240 X 256 matrix of amplitude values ranging from 0.0 to 1.0. In the data processing required to code these values for display, the amplitudes were quantized to 252 discrete levels from 4 to 255 (levels 0 to 3 were used to code the field surrounding the image). The function memory in the Comtal mapped this input matrix of amplitude levels to an output scale quantized to 8 to 64 levels with display coding in color, grey-scale, or some combination of color and intensities. The relationship between the input data amplitudes and the output coding was determined by the characteristics of the transfer function and the specific values assigned to threshold and gain. The principles underlying the definitions of these parameters are discussed below.

Figure 8 depicts the basic geometry relating threshold and gain to the transfer function. In this diagram the input amplitudes are scaled from 0.0 to 1.0 and are represented by the horizontal axis. The output amplitudes, represented by the vertical axis, are also scaled from 0.0 to 1.0 to correspond to the input, although any appropriate scale might be used. The output amplitudes are shown coded in six broad bands of color, plus white. A wide variety of color and grey scale coding schemes, ranging from 8 to 64 discriminable levels, might have been used for this diagram as described in a previous section.

In figure 8 the basic transfer function geometry is illustrated in terms of the "standard" coding (a). It is evident that with standard coding the full dynamic range of the input is mapped in a linear fashion to the full dynamic range of the output. As noted on the diagram, standard coding may be also defined as a threshold of 0.0 and a gain of 1.0. Other possible values of threshold and gain are shown on the diagram to indicate how the parameters are related.

Examination of the diagram reveals that the transfer function may be represented as the hypotenuse of a right triangle. The base of this triangle is equal to the percentage of the full range of the input amplitudes which is mapped by the transfer function, while the height of the triangle is the percentage of the full range of the output amplitudes which is actually coded. The threshold is seen to be represented by the intercept of the transfer function with the horizontal axis, while the gain is inversely related to the slope of the transfer function.

The geometry of the relationships illustrated in figure 8 may be summarized as follows:

Transfer function = the hypotenuse of a right triangle, where:

x = the percentage of the full range of the input amplitudes that is coded by the transfer function,

y = the percentage of the full range of the output amplitudes that is coded by the transfer function,

gain = 1/[slope of the transfer function (y/x)] = x/y, and

threshold = the x intercept of the transfer function.

Several numerical examples may clarify these definitions. Referring to figure 8(b), where the threshold is 0.0 and the gain is 0.5

gain = 1 / slope = 0.5/1.0 = 0.5.

Note that in this case the lower half of the full dynamic range of the input is mapped to the full dynamic range of the output. In figure 8(c), where the threshold is 0.5 and the gain is 0.5, the gain has the same slope as case (b) and is derived from the same values, except that

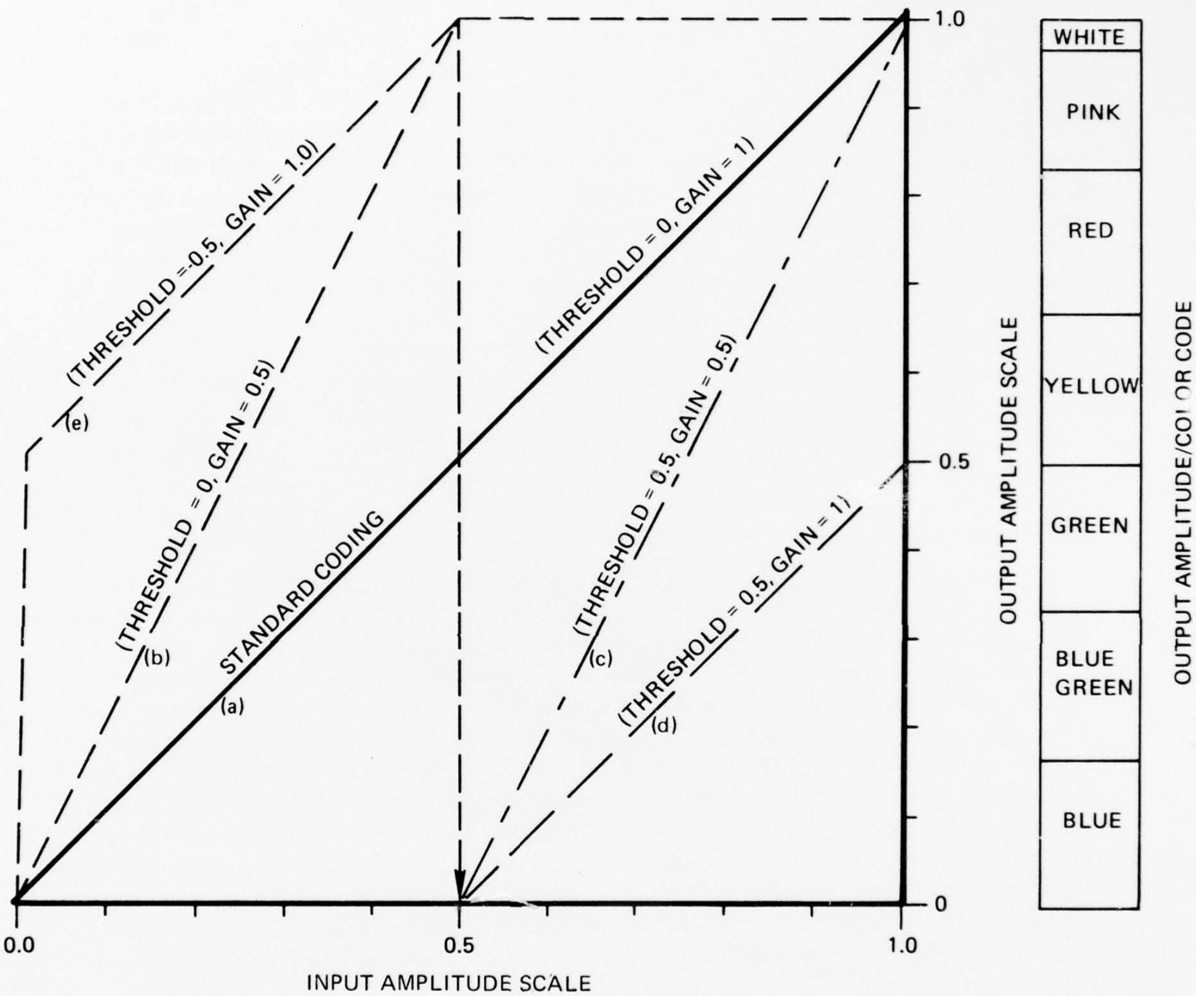


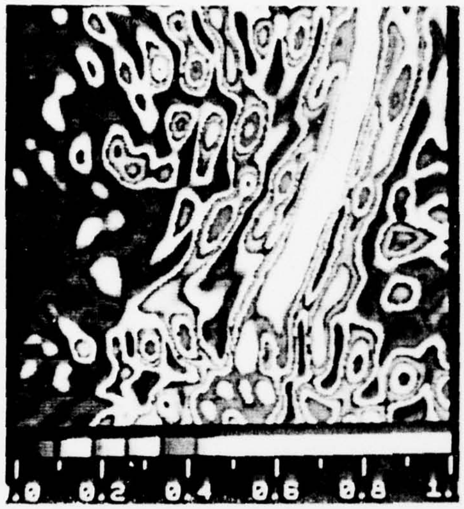
Figure 8. Input data to output color transfer functions illustrating threshold and gain parameters. The transfer function is represented as the hypotenuse of a right triangle whose base is the percentage of the full input range and whose height is the percentage of the full output range mapped by the transfer function. The threshold is represented as the intercept of the transfer function with the horizontal axis. Gain = 1/the slope of the transfer function.

in this case it is the upper half of the input which is mapped to the full range of the output. In figure 8(d), the relationships are somewhat different in that only the lower half of the dynamic range of the output is used to map the upper half of the output. Here, the gain is equal to 0.5/0.5, which is again equal to 1.0. The final case illustrated, figure 8(e), introduces the concept of a hypothetical negative threshold, in this case -0.5. Example (e) maps the lower half of the input to the upper half of the output, hence the gain of 1.0. Since the range from 0.0. to 1.0 is the full dynamic range of the input amplitudes, negative values of input amplitude cannot exist; however, for the computation of thresholds in Experiment 2, it was desirable to include those cases where the subject set the threshold below the standard coding of 0.0. For such cases, the threshold is defined as the x intercept of the transfer function when extended with the same slope to the horizontal axis to form a right triangle. One final relationship of those coding parameters should be noted. The algebraic sum of the threshold plus the gain defines the input amplitude mapped to the highest output coding, i.e., white in the examples in figure 8. This definition holds, however, only for values of threshold plus gain between 0.0 and 1.0. In example 8(e), the threshold of -0.5 plus the gain of 1.0 are equal to 0.5 which is seen to be the input value that is mapped to 1.0 or white. The importance of the threshold plus gain will become evident in later discussions of normalization and the results of Experiment 2. To further illustrate the coding principle discussed in this section, the effect of some representative settings of threshold and gain on image pattern features may be seen from the photographs in figures 9 and 10. Because it is a special case, a transfer function diagram has been included with figure 10(a) to illustrate how the gain of 1.5 is derived.

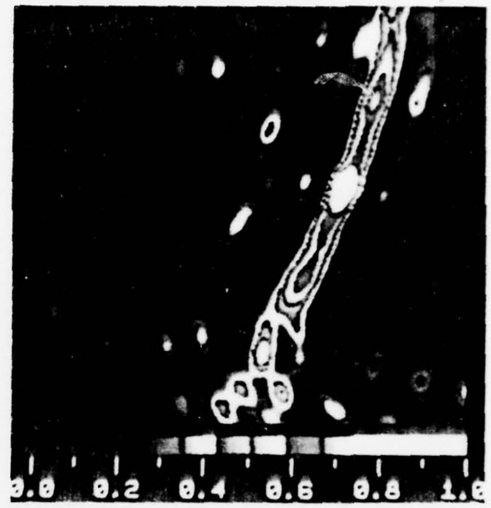
With this background understanding of figure 8, additional coding concepts are more easily explained. In the experiments reported here comparisons were often made between normalized and nonnormalized data images. A discussion of the autonormalization procedure used in these experiments is provided in this section. Autonormalization of the input data distribution is a process that divides each data amplitude by the highest amplitude in the data array. This normalizes the input data to a distribution of amplitudes ranging between 0.0 and 1.0. With the use of standard coding in the Comtal as illustrated in figure 8, it is evident that with these normalized amplitudes as input the full dynamic range of the output coding is matched to the normalized values from 0 to 1.0. This means that the highest amplitude is always coded white regardless of its amplitude prior to normalization. Note that autonormalization is equivalent to adjusting the gain so that the threshold plus gain is coincident with the highest amplitude in the input data. This relationship is used in Experiment 2 to relate the preferred settings of threshold and gain to autonormalization.

The significant difference between manual adjustment of threshold/gain and autonormalization is that in the latter case the input data are normalized to map one-to-one with the full dynamic range of the output scale. In the former case, the operator interactively adjusts the transfer function in the Comtal memory relative to the dynamic range of the input amplitudes. With manual adjustment, both threshold and the gain of the transfer function are under operator control Autonormalization adjusts the gain only (in a relative sense by renormalizing the data); the threshold remains constant at 0.0.

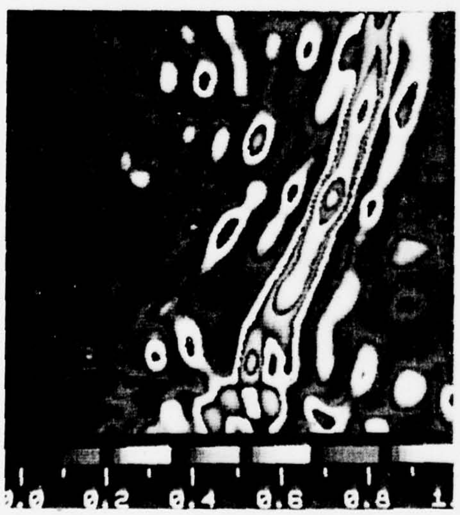
Before completing this section on coding parameters it should be noted that a greatly simplified model of the transfer function has been used in the discussion. For the multiband/one-step color coding schemes the transfer function was actually a stepped, linear ramp. Each step corresponded to a band in the output coding scale. The number of steps and the height of each step were determined by the gain and the number of discrimination levels in the output coding scale. In the multiband-multistep schemes the transfer function was again



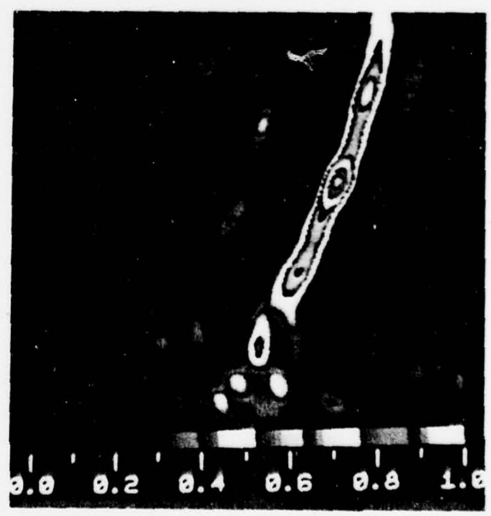
THRESHOLD: 0.0; GAIN: 0.5



THRESHOLD: 0.25; GAIN: 0.5



THRESHOLD: 0.0; GAIN: 1.0



THRESHOLD: 0.25; GAIN: 1.0

Figure 9. The effect of threshold and gain on image pattern features.

THRESHOLD: 0; GAIN: 1.5 THRESHOLD: 0.25; GAIN: 1.0

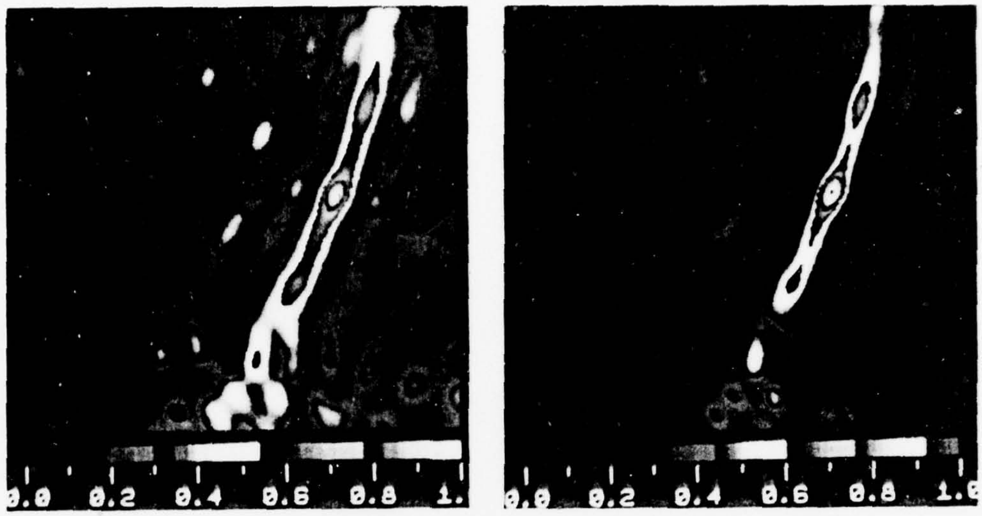
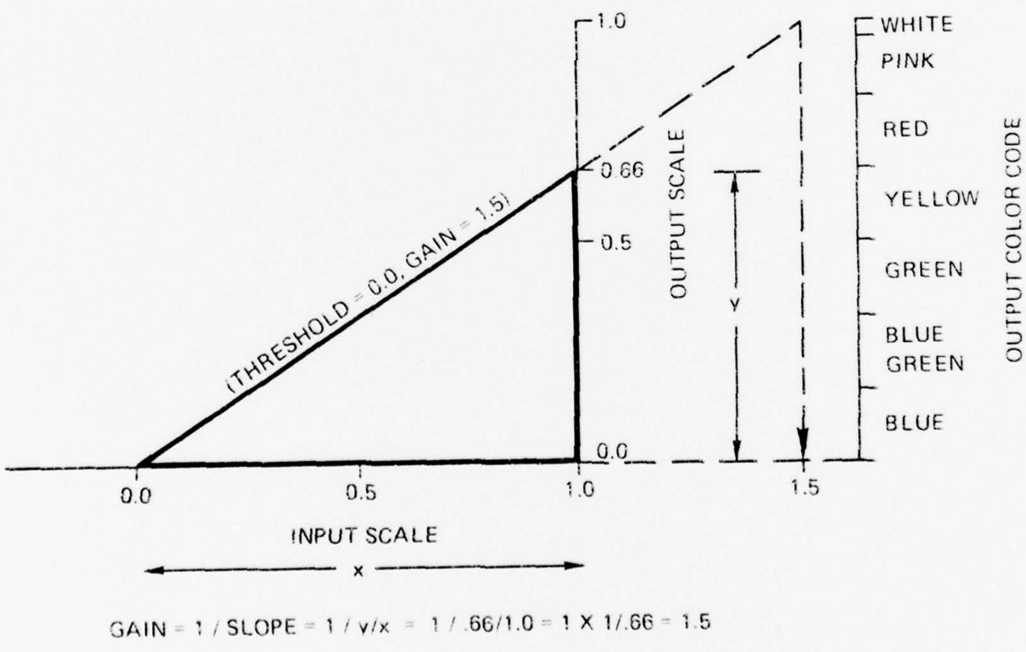


Figure 9. (Continued).



GAIN = 1 / SLOPE = 1 / y/x = 1 / .66/1.0 = 1 X 1/.66 = 1.5

Figure 10. Further illustration of the effect of threshold and gain on image pattern features. A transfer function diagram has been included to illustrate the special case of a gain value greater than 1.0.

a stepped linear ramp. However, in this latter case the steps corresponded to the intensity steps within each band.

The transfer functions for the grey-scale coding schemes were quite different. Because the response of the human eye to brightness differences is nonlinear, the transfer function must also be nonlinear to achieve equally discriminable output levels. Examples of the nonlinear transfer functions (quantization ramps) used for grey-scale coding in these experiments are in figure 5.

SUBJECTS

The experimental design for these experiments, known as a "subject by treatments" design, tested each subject under all of the conditions. In effect, each subject acted as his own control. Such a design is the most efficient test available for a relatively small, diverse sample of subjects. Eight or nine subjects were used in these experiments. The subjects were drawn from the professional staff of the NOSC Signal Processing and Display Division on the basis of their availability and their familiarity with the class of acoustic data used in the experiments. Actually, all but one of the subjects were members of the Display Branch within the division. All subjects had prior experience with the specific patterns used in the test and thus had prior knowledge of the significant pattern features. All had been exposed to the patterns at full signal strength and in the standard color code on many previous occasions. The subjects had little if any previous exposure to the patterns in the grey-scale. Thus, if the subjects brought any bias to the experiments it was in favor of color-coded images in the standard 6-band, 10-step format. With this potential bias present, any preference for grey-scale over color in these experiments assumed even greater significance.

EXPERIMENT 1: CODING SCHEMES

PROCEDURES

The purpose of Experiment 1 was to obtain ranking judgments on the perceptibility of six candidate coding schemes ranging from 8 to 64 levels of output amplitude coding. The rankings were made independently against four perceptibility criteria. The definition of the coding schemes, representing combinations of bands and steps, is described in detail on pages 6 through 26. Grey-scale and color quantization were included in Experiment 1 as parameters (i.e., variables held constant while the primary variables were changed to determine their effect). However, ranking judgments were made with respect to the coding schemes, only, in this initial experiment. Other parameters included were SNR at two levels and both normalized and nonnormalized data quantization. A detailed discussion of normalization was presented in Threshold and Gain Coding Adjustment, page 20.

Four separate criteria, listed below, were developed for this experiment to provide a more meaningful measure of the source of factors that might contribute to a judgment of overall perceptibility.

1. Sharpness of the demarcation (edge) between successive contours.
2. Overall contrast between adjacent contours (quantization levels).
3. Perceptibility of significant pattern features.
4. Perceptibility of equivalent amplitudes in widely separated areas throughout the image.

These criteria were selected on the basis of an informal assessment of the appearance of representative data patterns when coded with various schemes in grey-scale and color. An attempt was made to include criteria that would produce distinctly different rankings when applied to the various coding schemes. This approach can add significantly to our understanding of the basis for changes in the rankings from one condition to another.

The experimental design for Experiment 1 was represented by all combinations of the following variables: 6 (coding schemes) X 4 (criteria) X 2 (color/grey-scale) X 2 (SNR at +5.8 dB and –6.0 dB) X 2 (normalized and nonnormalized) X 9 subjects.

The test procedure involved the presentation of the six coding schemes in sequence on the Comtal in a self-paced or "paging" mode using the keyboard space bar. The subject was also able to call up any one of the six images for detailed comparisons by pressing specific letters on the keyboard. The subject paged through the six images for a given combination of the parametric variables, i.e., 2 (grey-scale/color) X 2 (SNR) X 2 (normalization) = 8 trials. His task was to rank the coding schemes from 1 to 6 against each of the four criteria in turn. The subject then initiated the next trial and repeated the ranking procedure. The same order of presentation for the parametric variables was followed for each subject to simplify the storage of data in the computer. This order is most easily described in programming terms. The most rapidly changing variable was the coding scheme. Then, in order, the criteria addressed by the subject: color or grey-scale, SNR, normalization, and finally the subject. Only one image pattern was used throughout. Representative image sets are shown on figures 2 and 4.

The instructions to the subjects, reproduced in full in appendix B, provide additional details on the test procedure.

RESULTS

The results of Experiment 1 are presented in figures 11 through 14.

Figure 11 presents the results in terms of the comparative average ranking of the six candidate coding schemes. The data are averaged over the nine subjects, the four ranking criteria, and three parametric variables (i.e., color-grey, SNR, and normalization). Subsequent figures break down the results in terms of various combinations of these parameters. Remembering that the coding schemes were ranked from 1 (highest) to 7 (lowest), figure 11 indicates that the overall ranking of the six coding schemes from highest to lowest was 3, 4, 1, 5, 6, and 2. The difference between schemes 3 and 4 is slight (not significant), as is the difference between 5 and 6. Scheme 1 is given approximately the same rank as 5 and 6. Figure 11 is useful mainly as an overall indication of comparative rankings, since the rankings change somewhat as a function of the criteria and the other parametric variables.

Figure 12 gives the results in terms of a color/grey comparison. Here we see a shift in color/grey ranking as a function of coding scheme, particularly for Scheme 4 and Scheme 1.

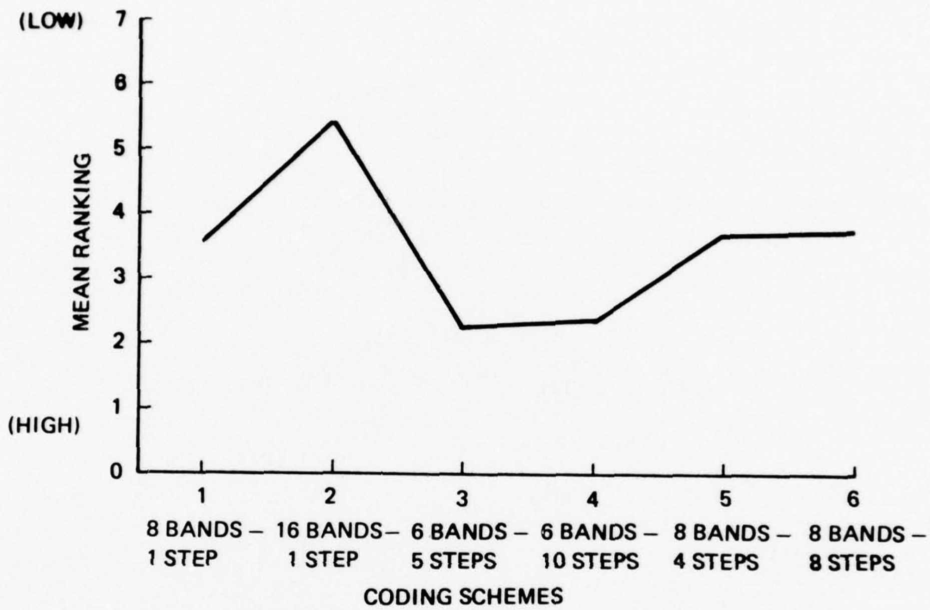


Figure 11. The mean ranking of the six coding schemes compared in Experiment 1.

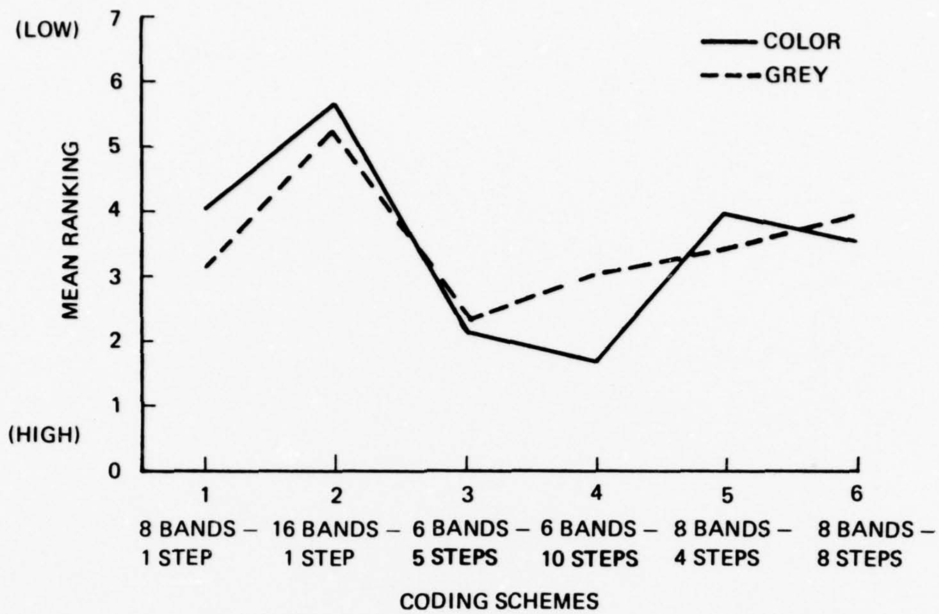


Figure 12. The mean ranking of the six coding schemes in grey-scale and color.

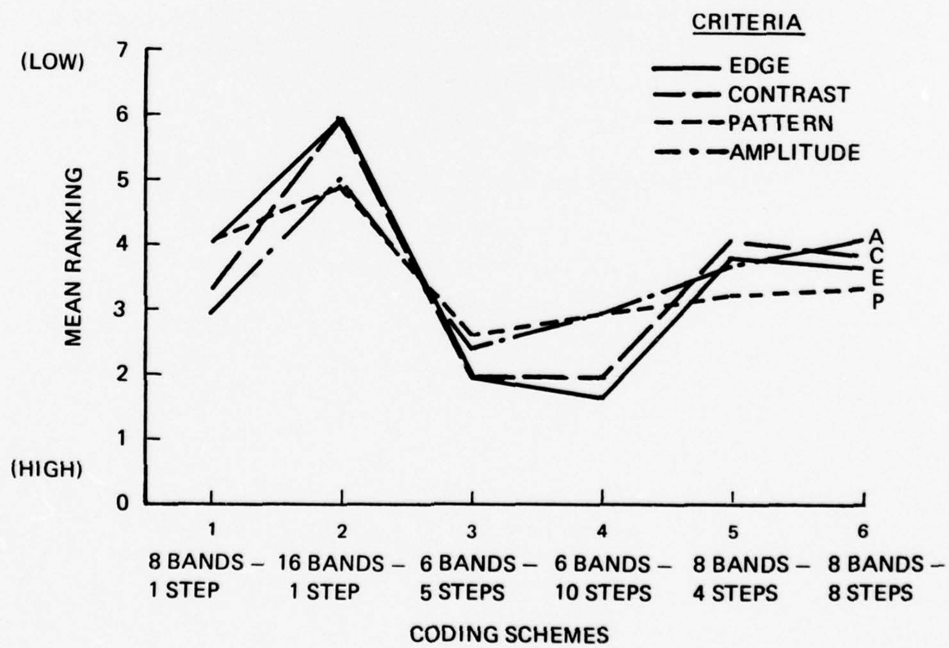


Figure 13. The mean ranking of the six coding schemes in grey-scale and color combined against four perceptibility criteria.

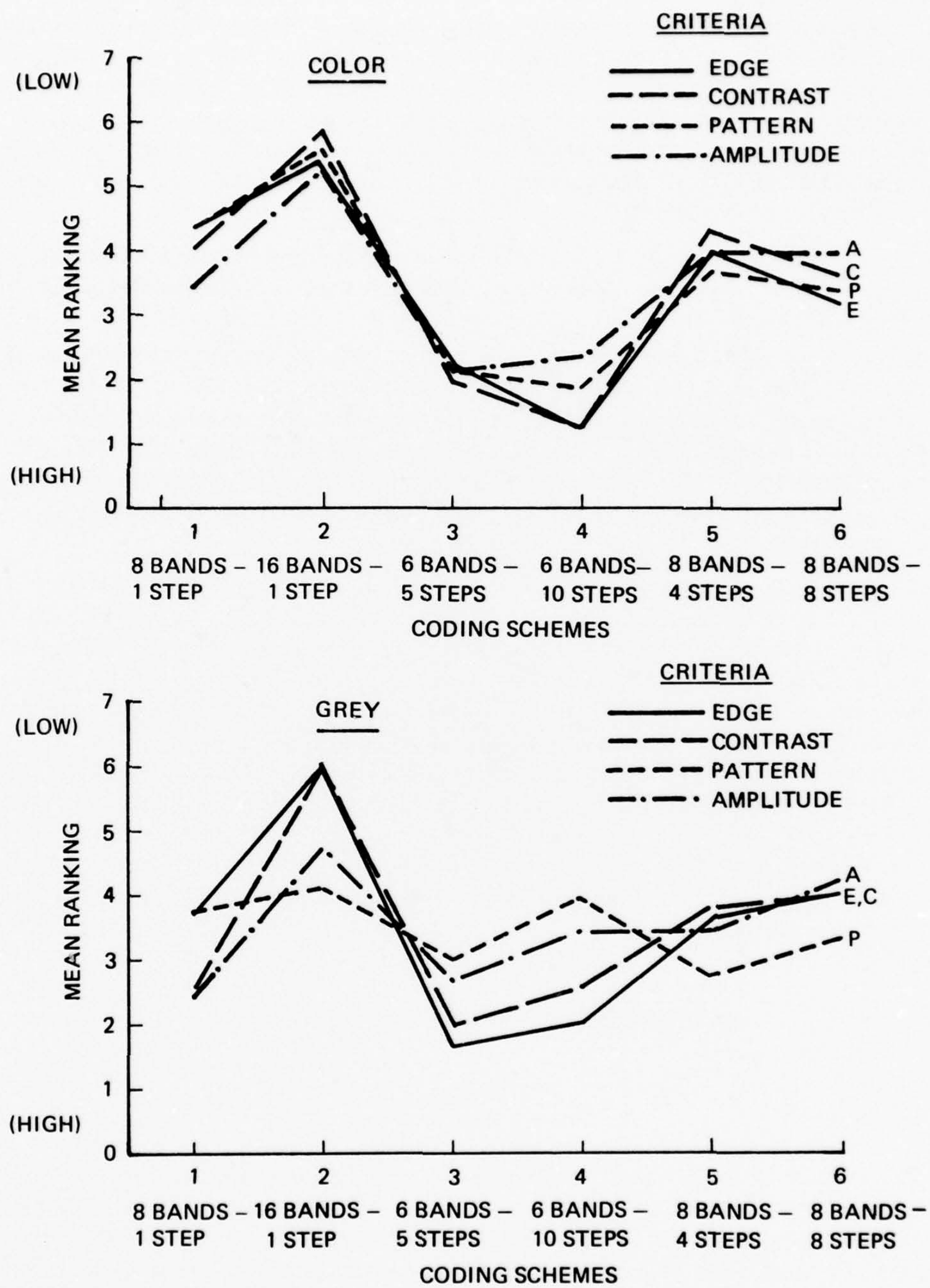


Figure 14. The mean ranking of the six coding schemes against four criteria and in grey-scale and color.

It is important to keep in mind, when reviewing these data, that Experiment 1 was not designed as a direct comparison of grey-scale and color coding. It would be misinterpretation of the results, for example, to conclude that grey-scale is preferred by one ranking unit over color coding with Scheme 1. The reason for the apparent difference may lie in the fact that when compared with some of the other grey-scale schemes Scheme 1 was noticeably better; whereas, the differences among the various schemes with color coding were not as marked. Results which directly compare the perceptibility of grey-scale and color codes are presented in Experiments 3 and 4.

The combined ranking in figure 11 shows schemes 3 and 4 to have approximately equal ranks. In figure 12 it is seen that for grey-scale Scheme 3 is preferred over Scheme 4; for color, the opposite is found.

Figure 13 presents the overall effect of the four different criteria as a function of coding scheme. These data indicate that edge and contrast generally produce similar rankings, as do pattern and amplitude. Figure 14(a) plots the mean ranking for the four criteria for the color-coded schemes. The differences in ranking, as a function of criteria, are seen to be slight, except for the pairing of edge and contrast and a similar pairing of pattern and amplitude on Scheme 4. However, with grey-scale coding, as shown in figure 14(b), the pairing of criteria is much more pronounced.

A detailed statistical analysis of these results, known as an Analysis of Variance (ANOVA), was computed using the UCLA BMD08V program (1971). An ANOVA of the significance of the differences between the means of the primary variables can be of considerable value when the differences are not as apparent as in this case. The ANOVA does, however, include a complete table of means, which greatly simplifies the documentation of the results. The ANOVA for these data confirms that the differences between the means for the combinations of variables, presented in figures 11 through 13, are statistically significant at the .001 level of confidence (i.e., a difference between means as large as this would occur less than one time in a thousand by chance alone). These significant differences include the coding schemes (figure 11), the interaction of color/grey with coding schemes (figure 12), and the interaction of criteria with coding schemes (figure 13). The interaction of criteria, coding schemes, and color/grey-scale (figure 14a, b) is significant at the .05 level of confidence.

EXPERIMENT 2: THRESHOLD AND GAIN

PROCEDURES

The purpose of Experiment 2 was to obtain preliminary information on the preferred manual settings of the threshold and gain quantization parameters, which may be adjusted by the operator on the Comtal display using the trackball. These preferred settings were needed for coding the test data sets for Experiment 3. A secondary objective of Experiment 2 was to compare the preferred manual settings of threshold and gain with those obtained by autonormalization.

A detailed discussion of the threshold and gain adjustment on the Comtal was provided on pages 6 through 26. The reader may wish to return to this section for a review of the major concepts and many procedural details which are not repeated here.

The primary independent variable in this experiment was the ratio of signal power to noise power (SNR) of the acoustic data images used as test material. Five levels of SNR ranging from +5.8 (or +2.1) dB to −11.4 dB were selected from the calibrated test data as the test images used. The construction of these data sets has already been described in detail. There were two data sets with different patterns in Experiment 2. Each SNR-pattern combination was presented three times. The images were coded only in color in the 6-band, 10-step scheme which was preferred by the subjects in Experiment 1. The experimental design for Experiment 2, therefore, consisted of all the possible combinations of 5 (SNR) X 2 (pattern) X 3 (replications) X 8 (subjects).

The test procedures in Experiment 2 were designed to simulate conditions where the operator might view a number of different patterns, with varying signal strengths, over a short period of time. Also, care was taken to avoid any bias introduced by a subject's previous setting on a similar pattern. For these reasons, a pseudorandom order of presentation of the test conditions was used to provide as much contrast as possible between successive trials. The presentation order alternated the two patterns, and one of the high SNR images was alternated with a low SNR image until the entire dual set of 10 images had been presented. The sequence was then repeated two more times.

By storing the entire sequence of images on tape it was possible to call up each image from the 1108 computer with a simple keyboard entry. The images were presented on the Comtal without annotation, including elimination of the usual calibrated color scale. The subject had no quantitative knowledge of his results – only the change to the image as he adjusted threshold and gain by means of the trackball. To insure an equivalent starting point for each trial, the image first appeared with standard nonnormalized coding. For example, if the highest data magnitude in the image were 0.3, the color coding of the peak in the data array would be just into the yellow band.

The task of the subject was the adjustment of threshold and gain by means of the trackball to make the significant pattern features of the image as distinctive as possible relative to the low level background. The objective, as stated in the instructions, was to improve upon the coding as first presented, if possible. As an aid to achieving this objective, the subject could return at any time to the coding of the image first presented by means of a keyboard entry. When satisfied with the threshold and gain settings, the subject sequenced to the next image by pressing the space bar.

Automatic recording and printout of the subject's threshold and gain settings were also provided. The measurement system recorded the starting point (threshold) and the range (gain) of the quantization ramp in the Comtal function memory in normalized units from 0.0 to 1.0 relative to the standard initial settings. If the subject's final setting of threshold was below the zero value of the input distribution, negative values were assigned. Illustrations of various combinations of threshold and gain are included in figure 8. Further details on the test procedures for Experiment 2 are provided in the instructions to the subjects, included in appendix B.

RESULTS

The results of Experiment 2 are shown in figure 15. The effect of the primary variable, SNR, is plotted in terms of the average threshold, gain, and threshold plus gain settings as described for the two patterns. The highest amplitude (signal or noise) in the

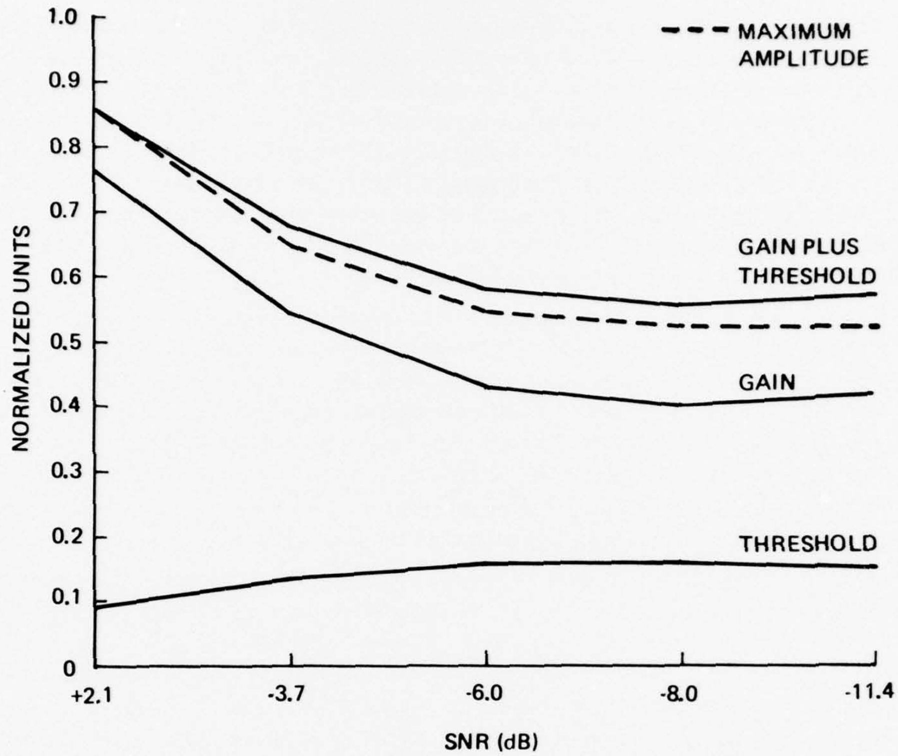


Figure 15. The mean threshold and gain settings in Experiment 2 as functions of SNR. The maximum amplitude (signal or noise) is indicated for comparison with the threshold plus gain function.

images at each SNR is also plotted for comparison as explained below. The parameter threshold plus gain always defines the amplitude in the input distribution, which is mapped to the highest level in the output color code. This is always white, if the value of the threshold plus gain is less than or equal to 1.0.

It may be recalled that autonormalization is accomplished by dividing the input data by the highest amplitude in the distribution (again, signal or noise). This renormalizes the data from 0.0 to 1.0 and the data are then mapped in the function memory to a color code in which the highest level is white, regardless of the data values prior to normalization. It is apparent, therefore, that if the value of the parameter threshold plus gain is coincident with the highest amplitude in the input data the subject has matched the coding for autonormalization.

The data presented in figure 15 (for the two patterns combined) show that the subjects do, in fact, closely match autonormalization, at least with respect to the coding for the highest amplitude. These data indicate that the subjects decreased their gain settings in a manner that closely matched the reduction in amplitude in the images. The threshold setting, on the other hand, remains relatively constant at 0.1 to 0.15 units of amplitude above the normalized 0 level. The preferred settings differ from autonormalization in that

the subjects raised the threshold a constant amount and, at the same time, decreased the gain proportionately to maintain the same coding for the highest amplitude (white).

Photographs of both the preferred coding and the nonnormalized coding for one of the patterns are provided in figures 16 and 17.

Detailed statistical analysis (ANOVA) for the data of Experiment 2 indicate, as might be expected, that the reduction in threshold plus gain values as a function of SNR is a true effect. The difference between the means is significant at a probability value less than the .001 level of confidence.

EXPERIMENT 3: GREY-SCALE VERSUS COLOR AT THE PATTERN THRESHOLD

PROCEDURES

The purpose of Experiment 3 was to determine the relative perceptibility of grey-scale and color-coded acoustic data images at the threshold of pattern perceptibility (i.e., at very low SNR levels where the patterning is easily confused with noise). In this experiment, relative perceptibility was defined in terms of a measured threshold value calibrated in conventional dB units. The threshold was determined by viewing a set of 11 images in which the amount of noise power was progressively increased relative to the signal power for a given combination of conditions. The development and use of data sets such as these is described in detail on pages 14 through 20.

The primary independent variables in this experiment were the grey-scale versus color coding and 11 SNR levels. Because the subject selected a single SNR level (the threshold) out of 11 possible choices, SNR is not treated as a primary variable in the analysis. Parametric variables (held constant on a given trial) were the two patterns used in Experiment 2 and three of the coding schemes used in Experiment 1. These were the 8-, 16-, and 60-level (6-band, 10-step) schemes. All data viewed in Experiment 3 were normalized to match the average threshold and gain settings at each SNR level preferred by the eight subjects according to the results of Experiment 2. Sample data sets in grey-scale and color are shown in figures 18 and 19.

The experimental design for Experiment 3 consisted of all combinations of the following variables: 2 (color and grey-scale) × 11 (SNR levels) × 2 (patterns) × 3 coding schemes) × 9 (subjects). This was also the order of presentation of the conditions from most rapidly changing to least rapidly changing. The same order of presentation for the combinations of parametric variables was followed for each subject.

The test procedure allowed the subject to view the 11 images on any given trial, four at a time on the Comtal display, as illustrated in figure 20. The subject was able to manually "page" forward or backward, one image at a time. By this means he could view the entire sequence of 11 images. He was also able to shift directly between the grey-scale and color versions of any given set of images on the display. A manually controlled cursor was available that automatically replicated its position at identical coordinates in the four images. This feature aided the subject in locating and tracking signal peaks on any sequence of four images as the signal progressively faded into the noise. Further details on the test procedures are described in the instructions to the subjects, included in appendix B.

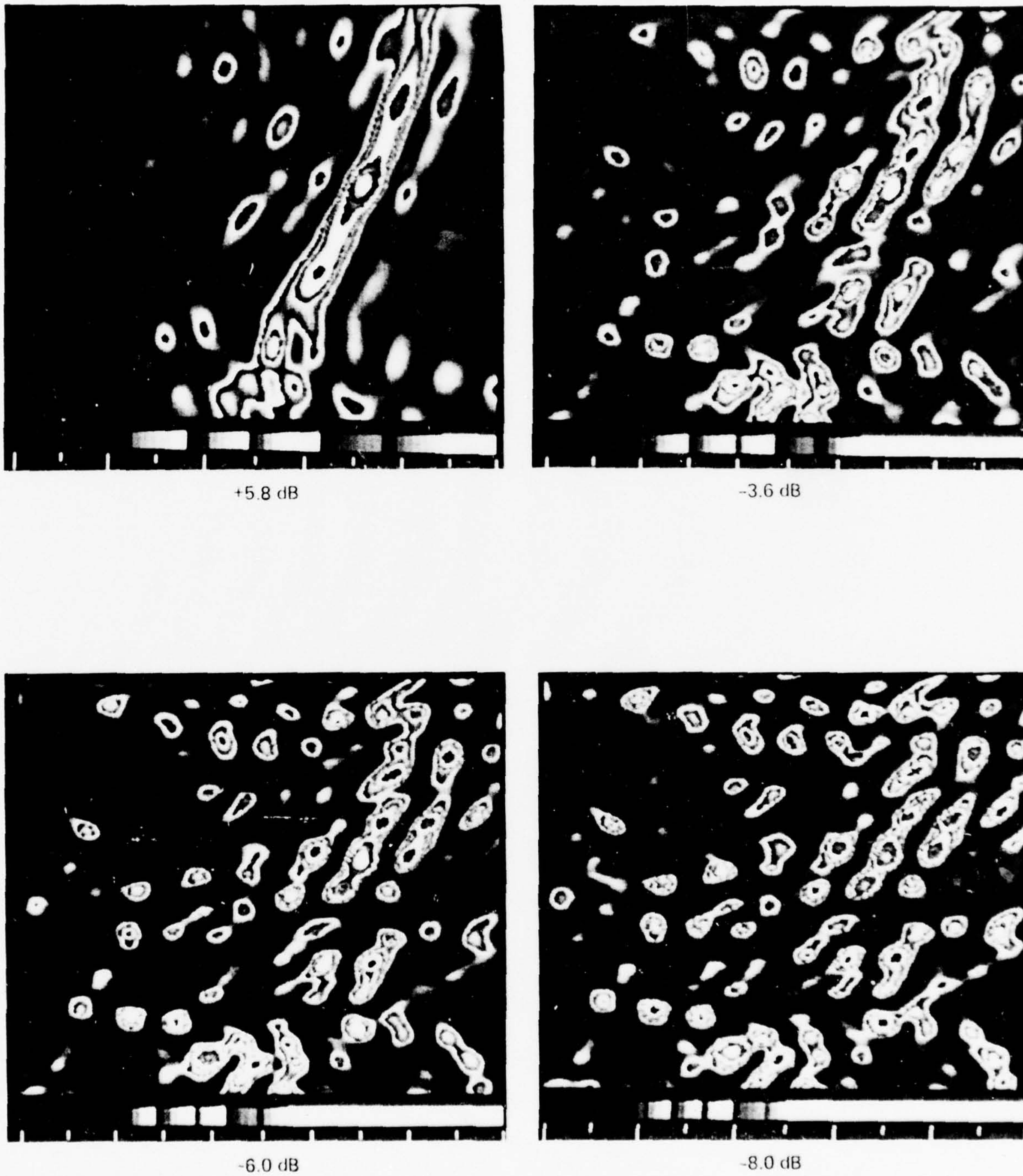
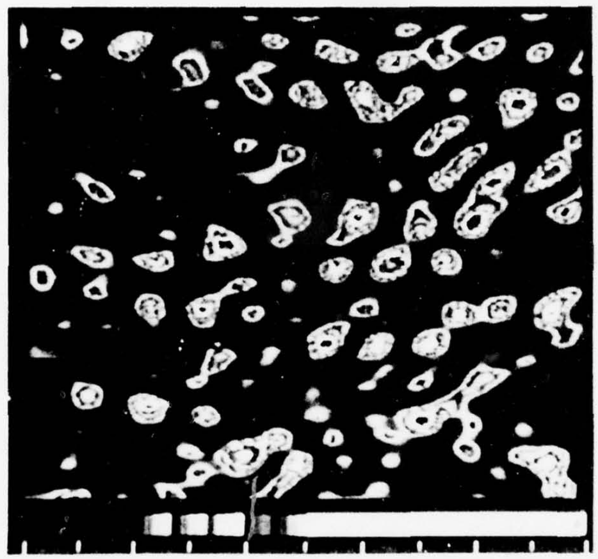
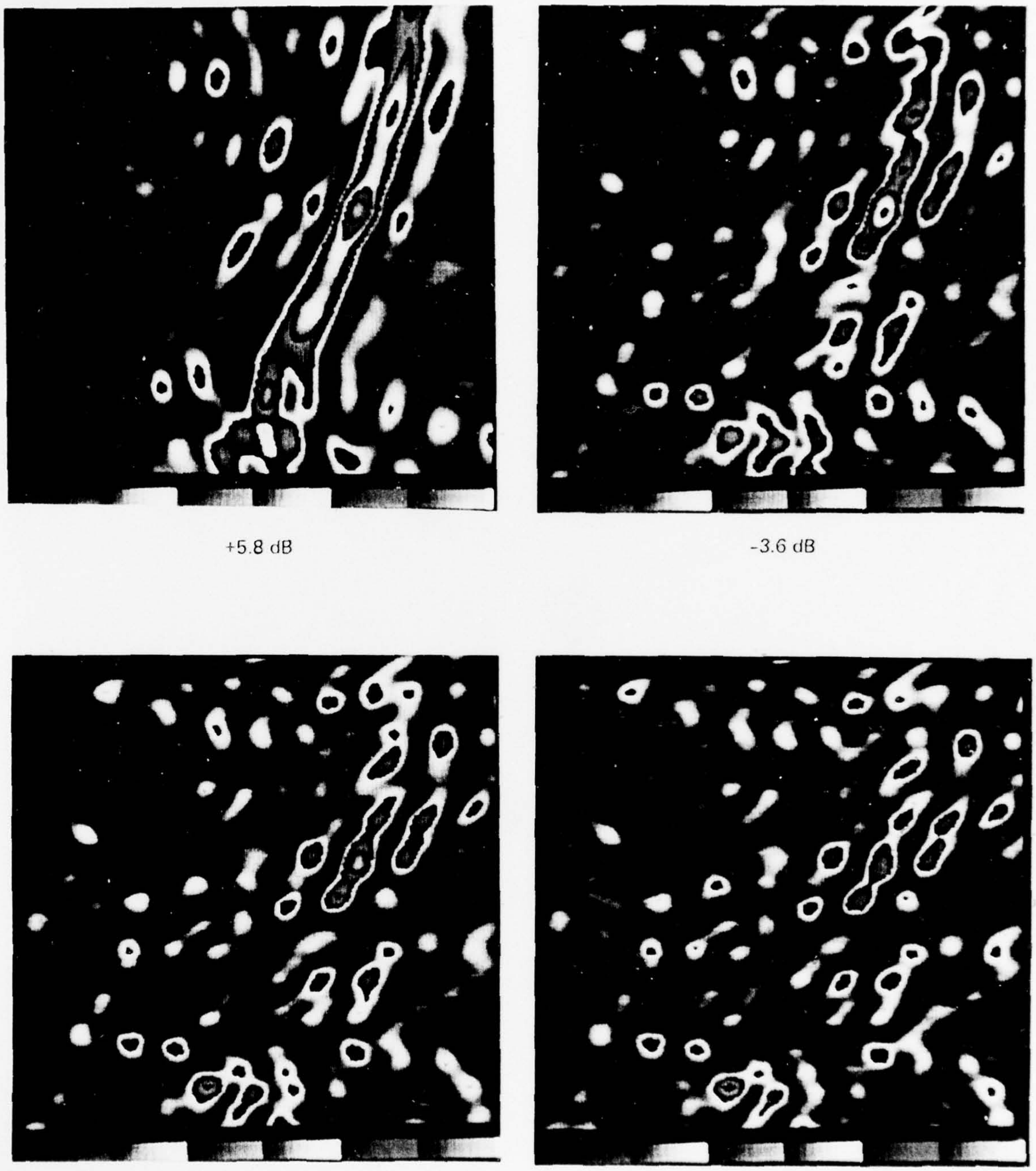


Figure 16. The preferred settings of threshold and gain illustrated by the image coding as a function of SNR.



-11.4 dB

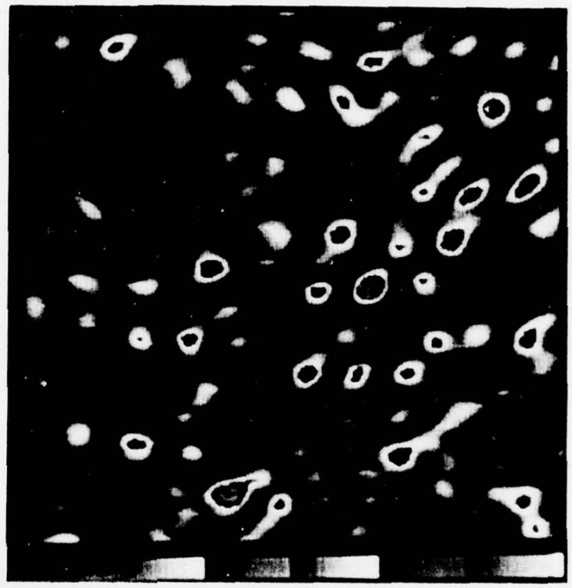
Figure 16. (Continued).



+5.8 dB   -3.6 dB

-6.0 dB   -8.0 dB

Figure 17. The initial nonnormalized image coding employed in Experiment 2 as a function of SNR.



-11.4 dB

Figure 17. (Continued).

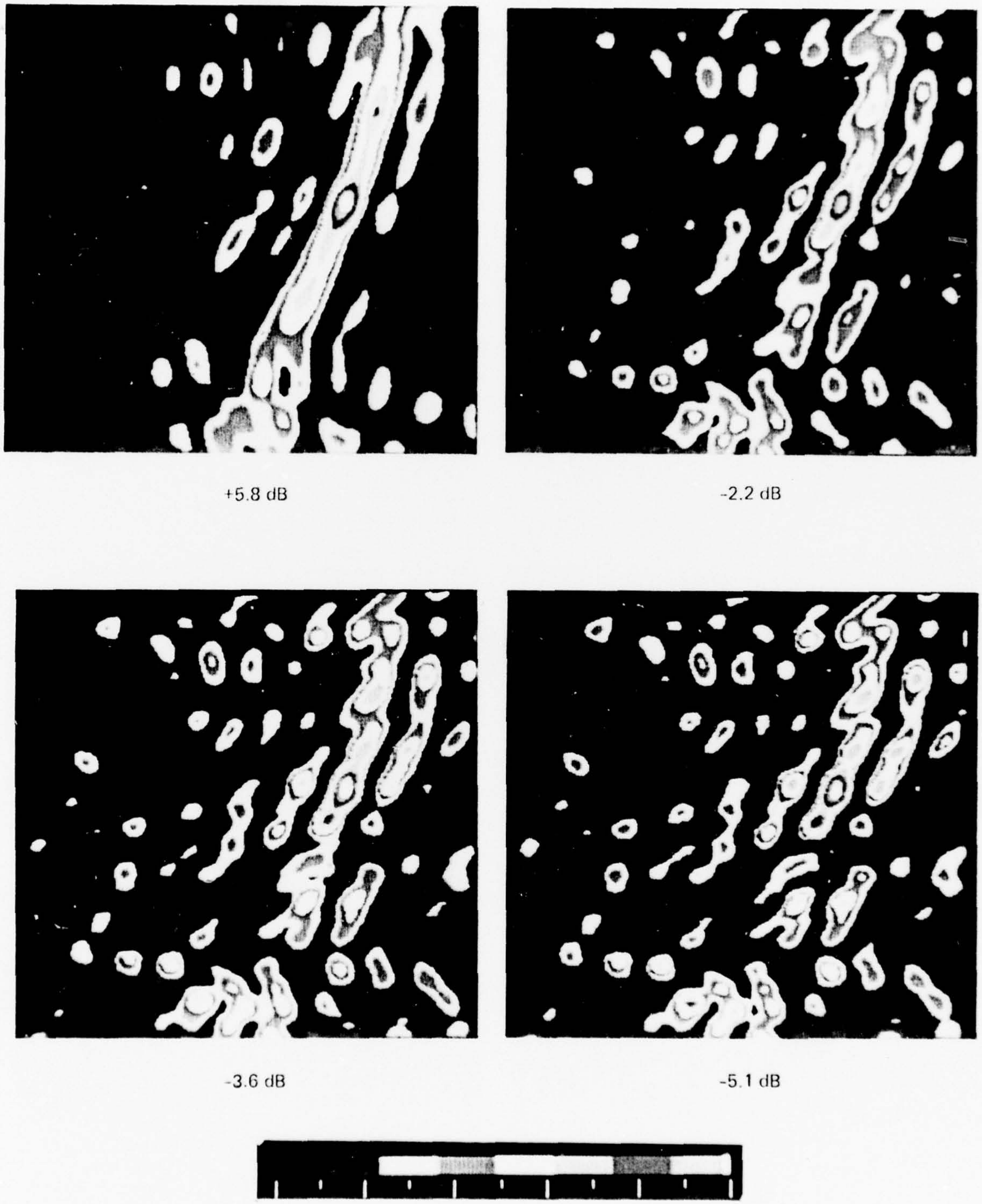
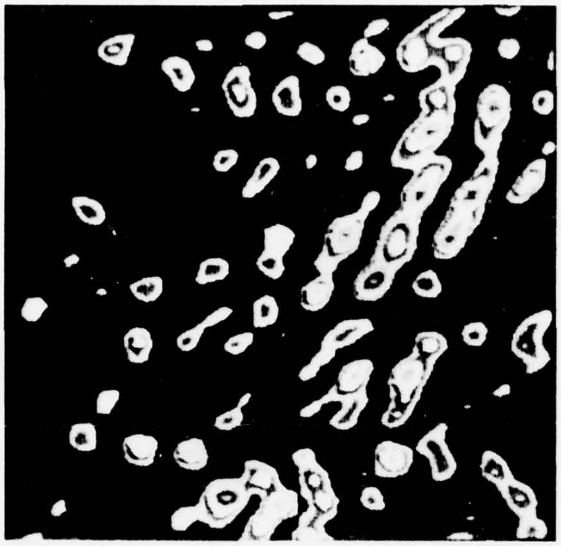
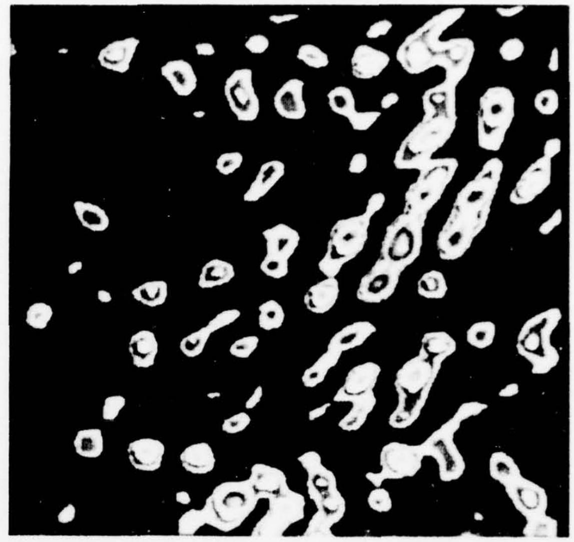


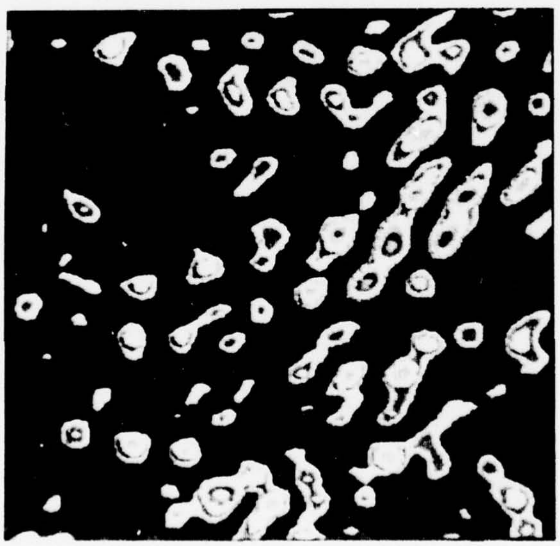
Figure 18. A test data set used for Experiment 3 with color coding.



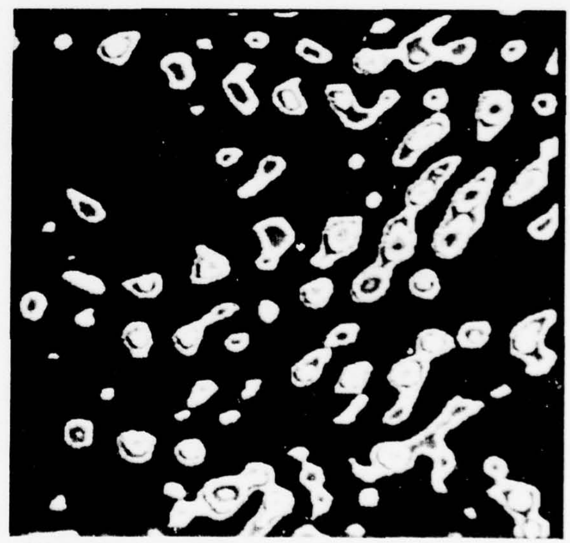
-6.0 dB



-6.9 dB



-8.0 dB



-9.3 dB

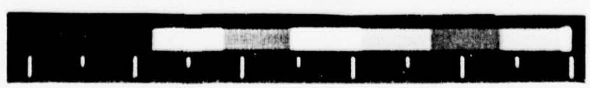
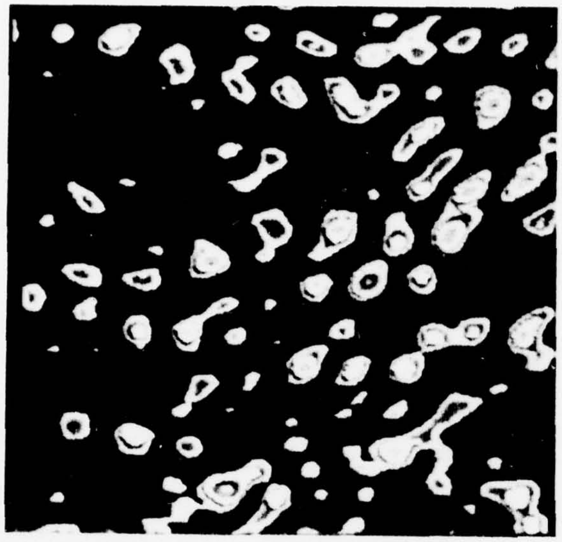
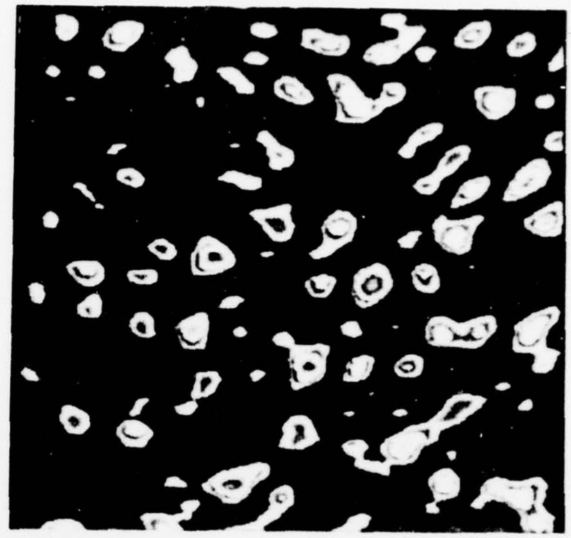


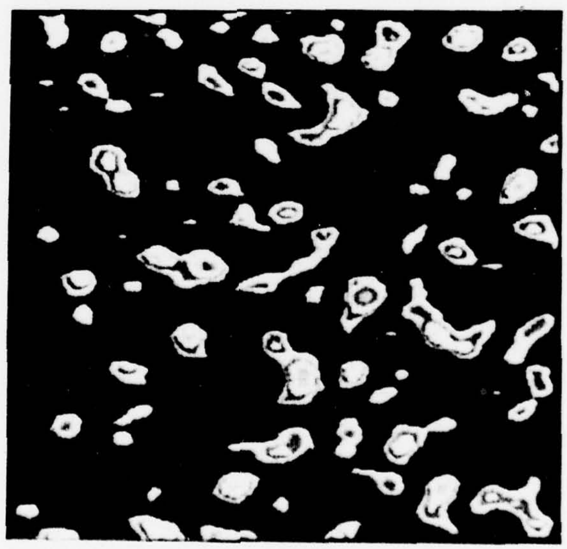
Figure 18. (Continued).



-11.4 dB



-15.5 dB



ALL NOISE

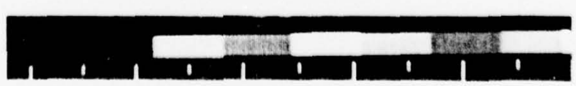


Figure 18. (Continued).

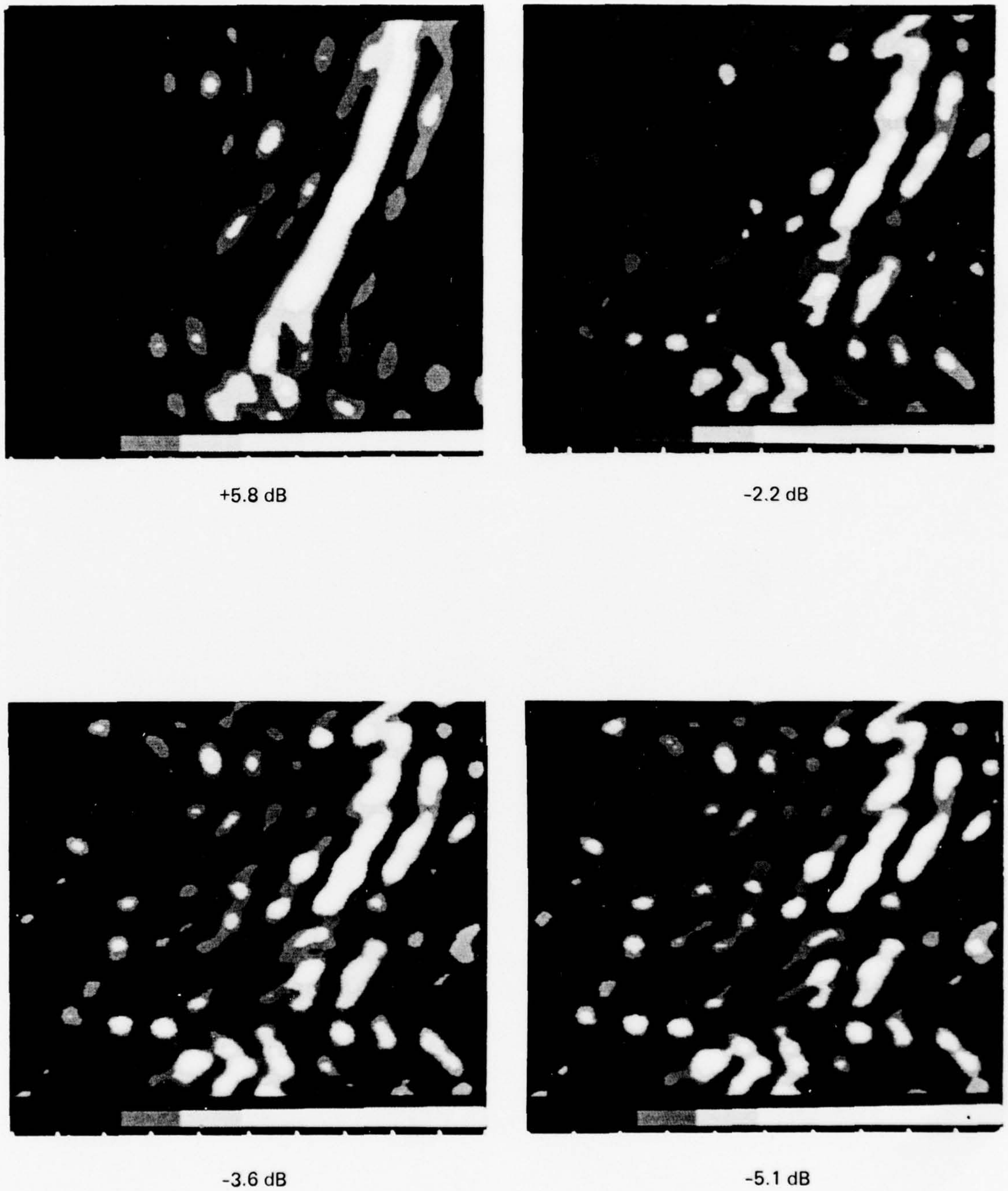


Figure 19. A test data set used for Experiment 3 with grey-scale coding.

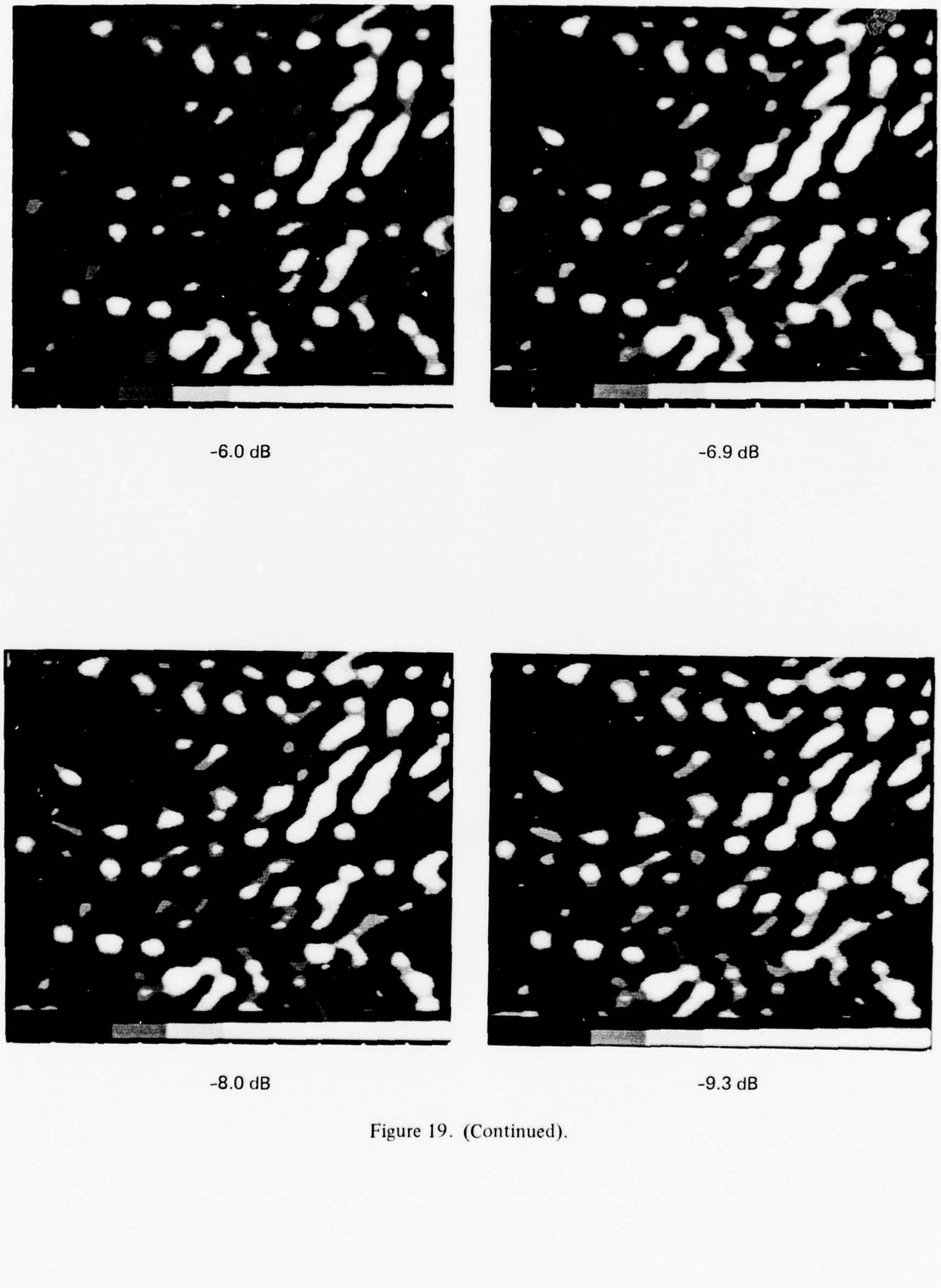
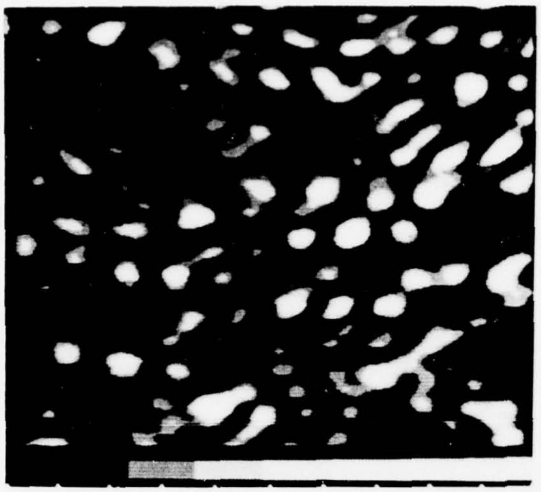
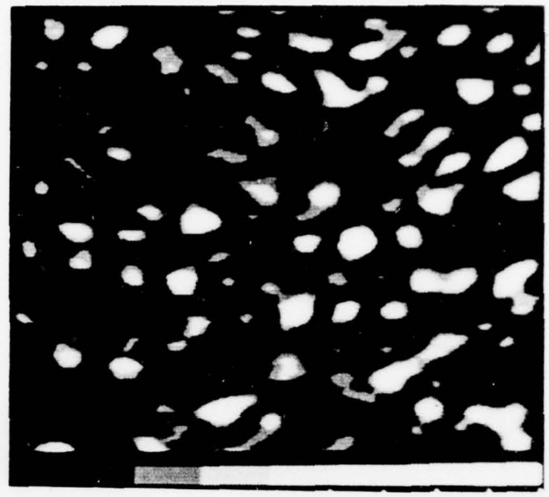


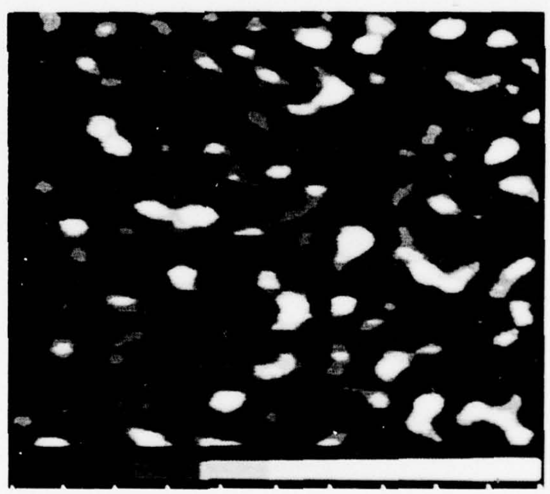
Figure 19. (Continued).



-11.4 dB



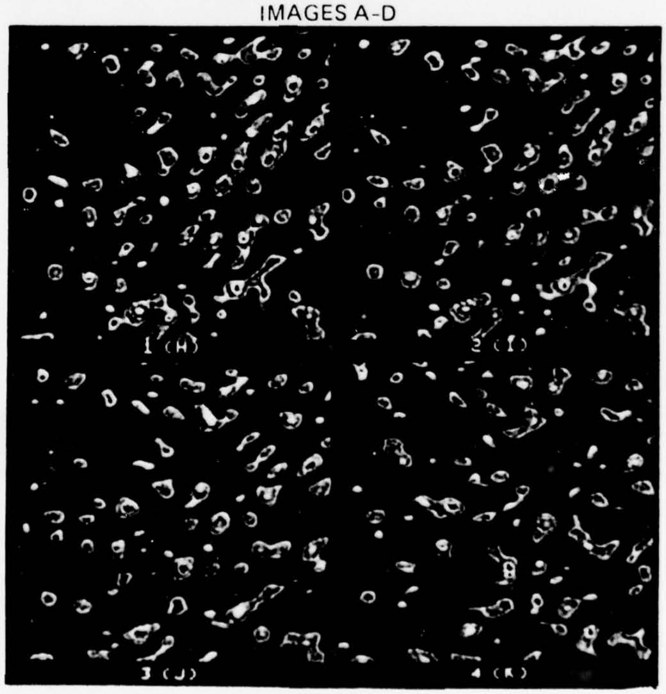
-15.4 dB



ALL NOISE

Figure 19. (Continued).

IMAGES A-D



IMAGES H-K

NOTE: THE REPLICATED CURSOR IS SEEN AS A SMALL WHITE DOT IN THE UPPER LEFT OF EACH IMAGE.

Figure 20. The four-image format used in Experiment 3. The subject paged forward or backward one image at a time or directly from A-D to H-K.

RESULTS

The results of Experiment 3 are presented in table 1. Table 1(a) gives the average threshold settings of the nine subjects for the grey-scale and color coding in arbitrary units from 1 through 11 corresponding to the 11 SNR levels from +5.8 dB to –30 dB. (For example, a threshold of 8.0 corresponds to an SNR level of –9.3.) Table 1 presents the average thresholds when the data for both patterns and the three coding schemes have been combined to arrive at a single mean value for grey-scale and for color. It is apparent that there is no difference between the grey-scale and color worth noting. Another way of presenting the results is by pattern for the grey-scale and color separately as in table 1(b). These data indicate that the color coding was slightly preferred on Pattern 1 while the grey-scale was the preferred coding on Pattern 2. Again, the conclusion is supported that overall there is no significant difference. Table 1(a) summarizes the data in still another way. The mean threshold values for the 8-, 16-, and 60-level coding schemes are shown separately for grey-scale and color coding. Again, there are no differences worthy of note.

Detailed statistical analysis (ANOVA) for Experiment 3, as expected, supports the conclusion of no difference between grey-scale and color coding under the conditions of this experiment (i.e., at the threshold of pattern perceptibility). The only significant difference found between means (.01 confidence level) is the interaction of the coding scheme with the pattern. This is merely statistical confirmation of the results shown in table 1(b), i.e., there may be a significant preference for grey-scale or color coding depending on the pattern viewed.

Table 1. Mean Threshold Determinations of Nine Subjects in Experiment 3 with Grey-Scale and Color Coding.

a. Average Threshold Setting

Subject	Grey	Color
1	7.83	7.83
2	8.50	8.33
3	8.50	8.00
4	8.50	8.33
5	8.00	8.00
6	8.00	8.00
7	8.00	7.67
8	8.17	7.83
9	8.33	8.67
Mean	8.20	8.07

b. Separate Results

	Grey	Color
Pattern 1	7.92	8.04
Pattern 2	8.48	8.11

c. Mean Threshold Values

	Grey	Color
8 Levels	8.1	7.9
16 Levels	8.3	8.3
60 Levels	8.2	8.0

EXPERIMENT 4: OVERALL RANKING OF GREY-SCALE VERSUS COLOR

PROCEDURES

The purpose of Experiment 4 was to obtain direct ranking judgments on the relative perceptibility of grey-scale and color-coded images at SNR levels clearly above the pattern threshold. The rankings were made independently against the same four perceptibility criteria used in Experiment 1. Three representative coding schemes used in Experiment 1 (i.e., the 8-, 16-, and 60-level (6-band, 10-step) schemes) were included as a parameter. Other parametric variables were SNR at three levels, normalized and nonnormalized quantization, and the two patterns used in Experiments 2 and 3.

The experimental design for Experiment 4 consisted of all combinations of the following variables: 2 (color and grey-scale) × 4 (criteria) × 3 (coding schemes) × 3 (SNR) × 2 (normalized and nonnormalized) × 2 (patterns) × 8 (subjects). This was also the order of presentation of the conditions, from most rapidly changing (grey-color) to least rapidly changing (subjects). The same order of presentation for the combinations of parameteric variables was followed for each subject. There were 36 trials (i.e., 3 (coding) × 3 (SNR) × 2 (normalization) × 2 (pattern)).

The task for the subjects required a direct comparison and comparative ranking of the perceptibility of the color and grey-scale versions for each image (trial) against the four perceptibility criteria. To facilitate this judgment the subject was able to shift between color and grey-scale as many times as he wished before making his decision. After rankings had been made against each of the criteria, the subject sequenced to the next trial and repeated the procedure.

The subjects were asked to make one of four judgments expressing the degree of difference in perceptibility of grey-scale and color on a scale of 1 (highest perceptibility) to 7 (lowest). These were: no difference (4-4), slight (3-5), marked (2-6), and extreme (1-7). The difference might favor either grey-scale or color. Further details of the test procedure are included in the instructions to the subjects in appendix B.

RESULTS

The results of Experiment 4 are presented in figures 21 through 24. The average ranking for grey-scale and color made by the eight subjects against the four criteria are shown in figure 21 in bar chart form. These data combine the results for the remaining parametric variables, i.e., SNR, normalization, and pattern. These data indicate that overall the coding schemes based on color were preferred over the grey-scale against each of the four criteria, as may be seen by the height of the "color" bars above the line indicating the ranking of 4-4 (no difference) for each criteria. The symmetry of the data for color and grey-scale results from the constraint placed on the ranking choices (i.e., 4-4, 3-5, 2-6, 1-7). Differences in the relative ranking as a function of the criteria are also evident from figure 21. The preference for color is greatest (2.5-5.5) for the contrast and amplitude criteria and least for pattern (3.4-4.5). It may be noted that the "perceptibility of significant pattern features" is the same criterion used in Experiment 3 to determine the threshold of pattern perceptibility

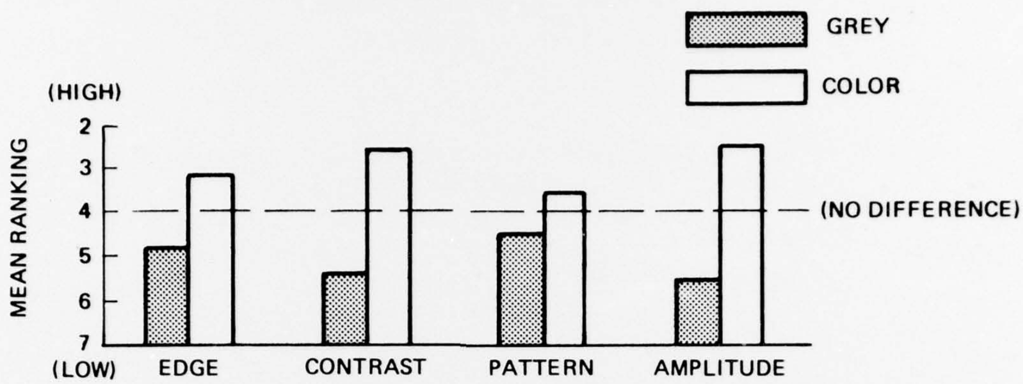


Figure 21. Mean ranking of the grey-scale and color coding against the four perceptibility criteria.

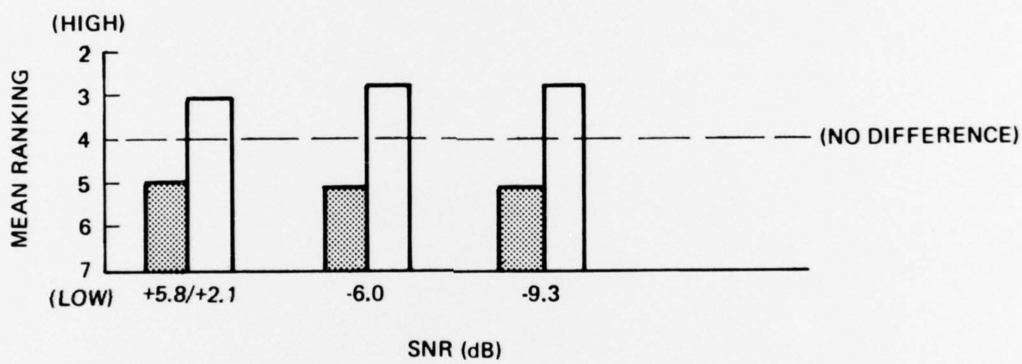


Figure 22. Mean ranking of grey-scale and color coding as a function of SNR.

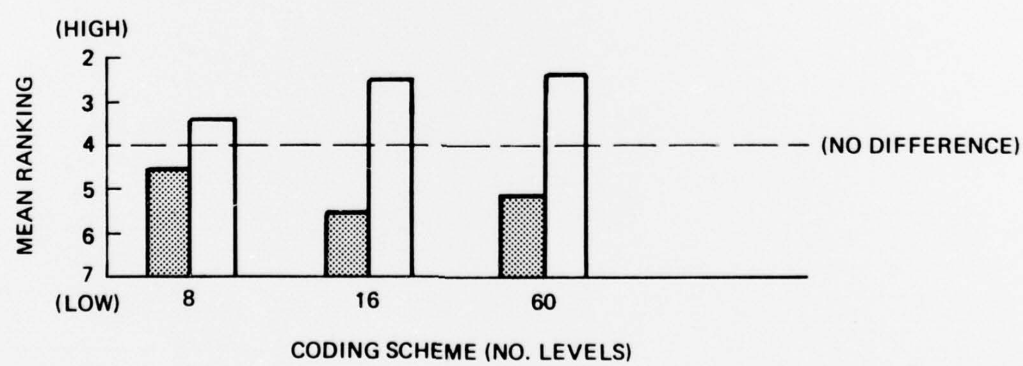


Figure 23. Mean ranking of grey-scale and color coding as a function of coding scheme.

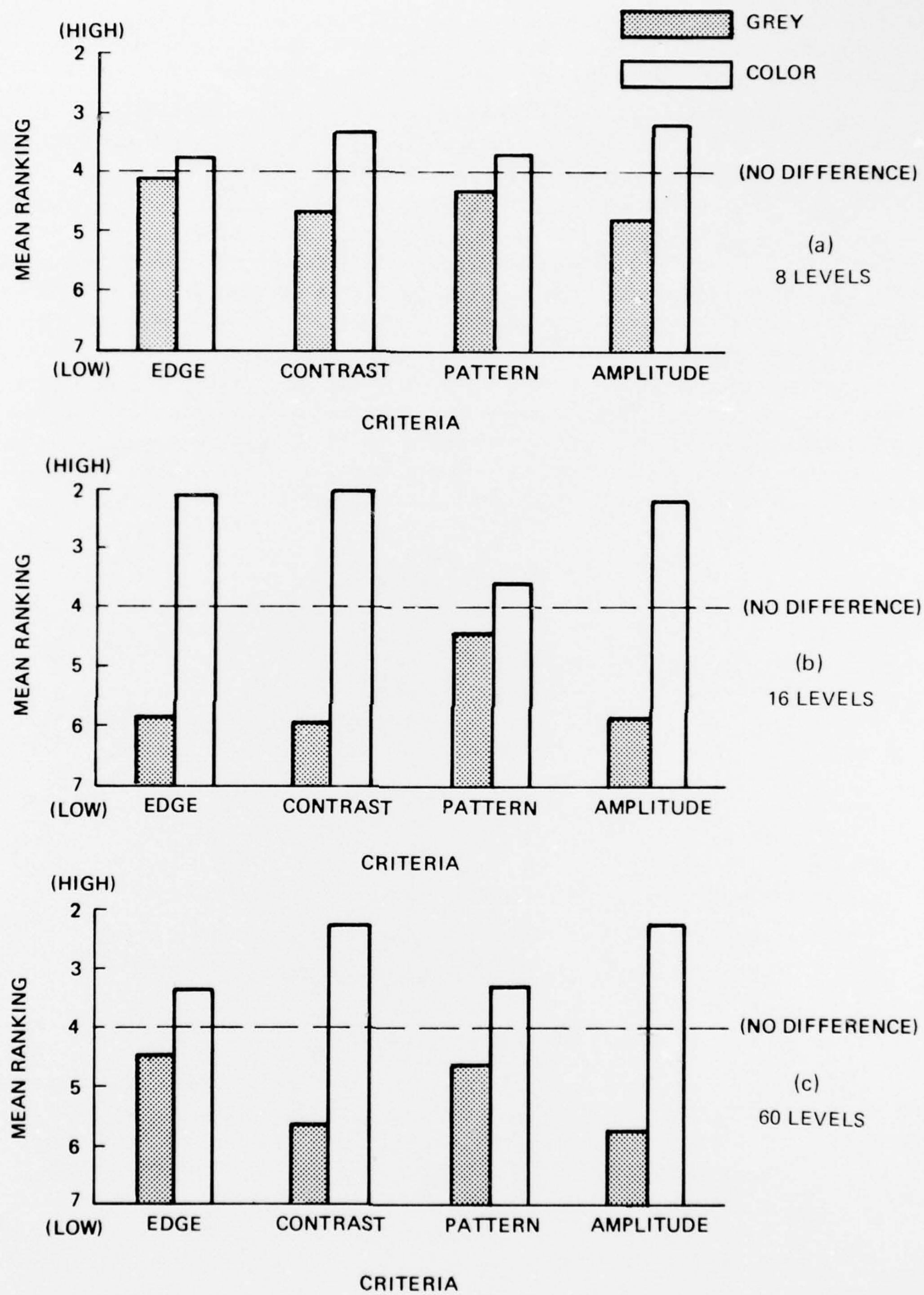


Figure 24. Mean ranking of grey-scale and color coding as a function of the four criteria and the three coding schemes.

for grey-scale and color. Thus, the results are consistent with those of Experiment 3, keeping in mind that the lowest SNR level used in Experiment 4 (−6.0 dB) was considerably above the average threshold obtained in Experiment 3 (−9.3 dB).

Figure 22 indicates that there is little change in the overall ranking as a function of SNR level from +5.8/+2.1 dB to −9.3 dB. From figure 23 it is seen that difference in ranking is least for the 8-level coding scheme and greatest for the 16-level scheme.

Figure 24(a), (b), and (c) break the ranking data down in terms of both the criteria used and the three coding schemes. The relatively small difference in the case of the pattern criterion noted for figure 21 is carried over in figure 24. The difference is most pronounced for the 16-level scheme. The differences for the other criteria are also the most extreme for the 16-level scheme. Note also that the smallest differences are found for the 8-level scheme, particularly for the edge and pattern criteria.

Detailed statistical analysis (ANOVA) for these data confirms that the differences noted are true effects with a .001 confidence level or better. Significant differences between means include the basic color/grey-scale comparison, the interaction of the criteria with the color/grey (figure 21), the interaction of coding scheme with color/grey (figure 23), and the interaction of coding scheme, criteria, and color/grey (figure 24(a), (b), and (c)).

CONCLUSIONS

SUMMARY OF RESULTS

In summary, the major results and conclusions from the experiments conducted under this program are presented in the following paragraphs.

EXPERIMENT 1

The ranking of the overall perceptibility of six candidate coding schemes was

1. (Highest) 6 bands, 5 or 10 steps per band (30-60 discriminable levels)
2. (Medium) 8 bands, 1 step per band (8 levels); 8 bands, 4 or 8 steps (32 to 64 levels)
3. (Lowest) 16 bands, 1 step per band (16 levels)

With color coding, this ranking order was the same when measured against four different perceptibility criteria. With grey-scale coding, a similar ranking was found for edge and contrast criteria. However, the perceptibility of significant pattern features and equivalent amplitudes produced little difference in ranking as a function of the coding scheme.

EXPERIMENT 2

The preferred manual settings of threshold and gain approximately matched the full dynamic range of the output to the actual range of the input data, except that the threshold was raised to exclude the lowest 10 percent of the input data amplitude range.

EXPERIMENT 3

No difference was found in the threshold of pattern perceptibility for grey-scale and color coding for three representative coding schemes.

EXPERIMENT 4

Overall, for above threshold patterns, color coding was preferred over grey-scale when compared directly against four perceptibility criteria. This preference was greatest for the perception of equivalent amplitudes and least for the perception of significant pattern features.

CONCLUSIONS

For the class of acoustic data images investigated the following conclusions apply.

COLOR VERSUS GREY-SCALE

A color display has some advantages over black and white, if the objective is the perception of patterning, contours, and equivalent amplitudes. If the objective is solely discriminating patterns of very weak signals in noise, the color display has little advantage over black and white.

LEVELS OF QUANTIZATION

A display system capable of 32 to 64 levels of grey or color quantization has advantages over an 8- or 16-level system for overall perceptibility of patterning, contours, and equivalent amplitudes. Note, however, that the advantage of more than 8 discriminable levels lies in the use of band-step coding. In the preferred 6-band scheme with 60 total discriminable levels, the 6 color bands divide the amplitude range into 6 levels; the 10 intensity steps produce an intensity gradient within each band that highlights the contours and allows finer amplitude discrimination.

Eight discriminable level displays were ranked very nearly as high as displays with 32 to 64 levels. Unless there is an overriding need for the greater dynamic range, an 8-level display system appears to be quite adequate as a second choice after the 64-level system.

CODING SCHEME

If from 32 to 64 levels of output coding are available, band-step coding is strongly recommended over a simple unidimensional quantization ramp. No more than 6 bands with either 5 or 10 intensity steps are required for good perception of patterning and equivalent amplitudes.

If from 8 to 16 levels of output coding are available, the preferred quantization scheme is 8 easily discriminable levels. Sixteen-level coding is not recommended for pattern perception because there are too many levels for the necessary discrimination at the critical higher amplitude intensities and too few levels for effective band-step quantization.

INTERACTIVE THRESHOLD/GAIN ADJUSTMENT

The capability for interactive adjustment of the threshold and gain parameters does not appear to improve perception of pattern features and does not appear to be necessary for a basic system displaying the types of acoustic data images considered in these experiments. Quantization that matches the preferred settings of threshold and gain (determined in Experiment 2) can be approximated through a simple software routine which normalizes the input data so that the full output dynamic range is utilized and which adjusts the threshold to exclude the lowest 10 percent of the input data amplitude range.

REFERENCES

1. Butler, W. B., "Color and Black and White Display of Sonar Information," Tracor Document T71-AU-9621-U, 25 February 1971.

2. Butler, W. B., and McKemie, W. M., "Engineering Guidelines for the Use of Color in a Sonar Display," Tracor Document T73-AU-9550-U, 1 April 1974.

3. Christ, R. E., and Corso, G. M., "Color Research for Visual Displays," ONR Report No. ONR-CR213-102-3, July 1975.

4. Lamar, D. L., Merifield, P. M., et al., "Evaluation of Pseudocolor Transformations of Skylab and Landsat Images and Test Charts, Part V, Final Report, SKYLAB EREP463 Results," CalESCO Technical Report 76-1, 1976.

APPENDIX A
THE COMTAL DISPLAY SYSTEM

All of the experiments described here were conducted in the NOSC Surveillance Display and Image Processing Laboratory. The test data used to compare grey-scale and color coding were viewed directly on the Comtal 8300 Display System pictured in figure A1, center, back. This display system, which is the central feature of the laboratory, provides the user with the capability to display and interactively manipulate high resolution color images. The host computer for the laboratory is the UNIVAC 1108. The 1108 computer and the Display and Image Processing Laboratory are both located in a secure area and can be used for processing classified data. The COMTAL accepts I/O commands and data from the host 1108 computer and internally performs scan conversion and refresh storage operations on the data to produce bright flicker-free displays in grey-scale or color. Pertinent features of the COMTAL system include:

- 512 X 512 picture element (pixel) resolution with a 1:1 aspect ratio
- Selectable data display modes including grey-scale, pseudocolor, and true color
- Three independent images stored at one time
- Eight bits of intensity coding stored for each pixel (six bits displayed)
- Scan conversion for moving (rising or falling) raster presentations
- Graphics overlay for superimposing outlines, grids, or alphanumerics on displayed images
- Trackball positioning of a target pointer using a trackball input control device.

Other display equipment pictured in figure A1 includes a digital-to-video image conversion system for color video-tape cassettes (extreme left, back); and ADVENT video-beam system (extreme left, front) which projects color video images onto a 7-foot-diagonal screen; and a STATOS V electrostatic printer/plotter (right, back).

Extensive user-oriented software has been developed in support of Display and Image Processing Laboratory activities. An operating system complete with a library of FORTRAN callable user routines has been written for the Comtal 8300 system. Using this system, highly interactive operations involving data retrieval and editing, alphanumeric and graphic overlay generation, trackball target control, and image manipulation using function memory modification can be performed on demand terminals located in the laboratory area. In addition, many user application routines exist including algorithms for transform and spatial domain image processing, continuous surface graphics, automatic scoring and analysis of experiments, image formatting for computer processed and natural images, stereoscopic imagery, and high refresh rate (up to 4 frames/second) loading of image sequences.

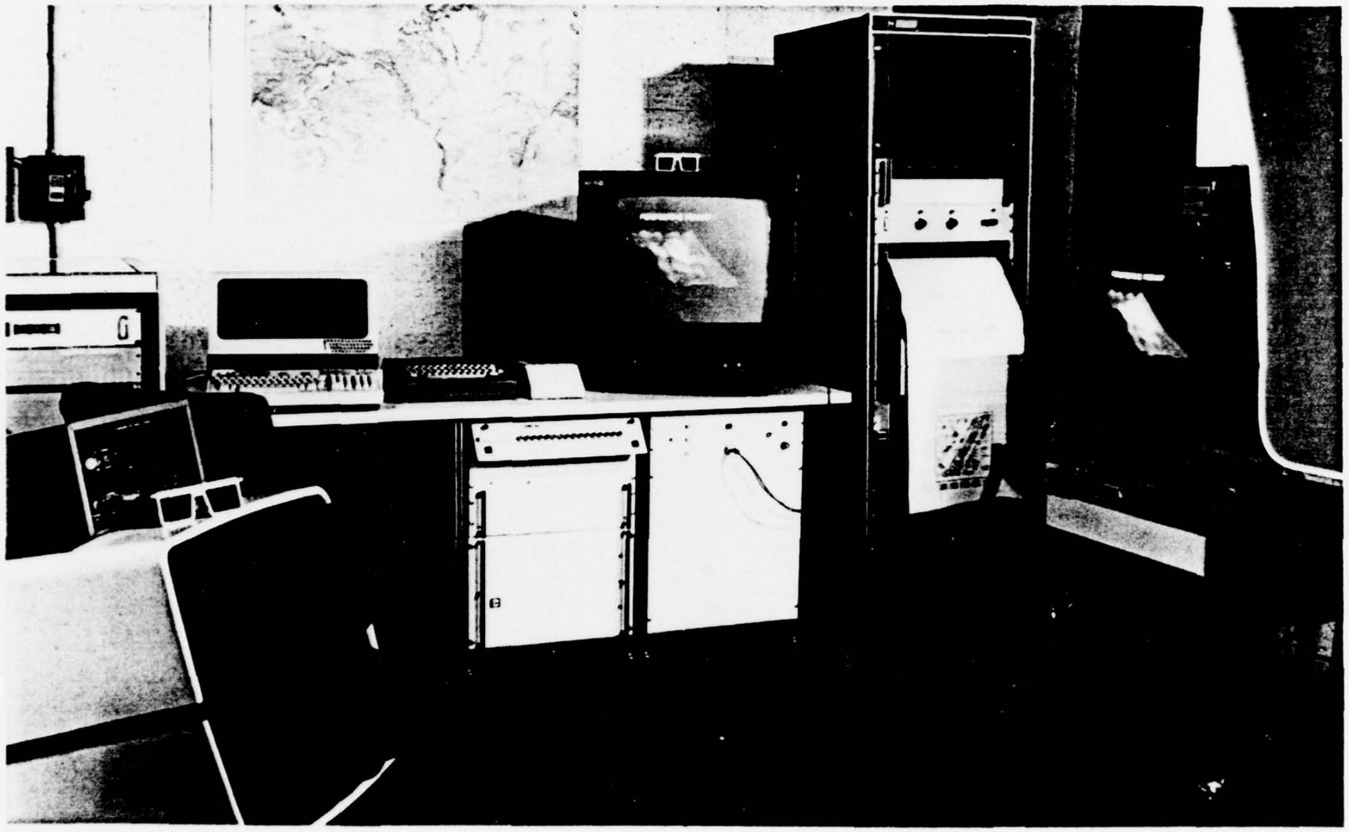


Figure A1. NOSC display and image processing laboratory.

APPENDIX B
INSTRUCTIONS TO THE SUBJECTS

EXPERIMENT 1: CODING SCHEMES

"This is a test of the relative perceptibility of various coding schemes. You will view six schemes (the primary variable) which have been selected to cover the range from 8 to 64 levels of quantization. These schemes will be presented in both color and grey-scale. In addition, there will be two other parametric variables. There will also be two SNR levels, and coding will be normalized and nonnormalized. The experimental design is 6 (schemes) X 2 (color/grey) X 2 (SNR) X 2 (normalization) for a total of 48 image conditions. The six coding schemes will be presented in a standard sequence on the Comtal for each of the 8 combinations of parametric variables (2 X 2 X 2).

"Your task is to compare the perceptibility of the six schemes against criteria listed below and to decide upon a ranking from 1 (highest perceptibility) to 6 (lowest perceptibility). A response sheet has been provided for recording your rankings. The six images will appear on the Comtal in a standard order as explained below, and you are free to examine each image as long as you wish. When you press the space bar you will advance to the next image in the sequence. You may do this at any time during the test. After your initial examination of all six schemes, you will be able to recall any one of the images as many times as you wish for more careful examination and comparison by pressing one of the corresponding keys from 'a' to 'f'. This direct comparison procedure should be the basis for your final ranking judgments. When you have recorded your rankings, you can advance to the next condition by pressing the 'Return' bar.

"A set of photos for the six coding schemes for color and grey-scale mounted in the order of presentation (i.e., A to F) is provided to aid you in recording your rankings. Refer to the photos for orientation only. Note from the scales that the sequence is in the following order: 8 bands, 1 step; 16 bands, 1 step; 6 bands, 5 and 10 steps; and 8 bands, 4 and 8 steps.

"This is primarily a test of the relative perceptibility of the six coding schemes. The other variables are included to test for a possible change in rankings as a function of the parameter values, i.e., interaction effects. To aid in the interpretation of any interaction effects you will make independent sets of rankings against the following four criteria:

a. Sharpness of the edge between successive contours
b. Overall contrast between adjacent contours
c. Perceptibility of significant pattern features
d. Perceptibility of equivalent amplitudes in widely separated areas throughout the image.

"It is quite likely that the rankings will be different for different criteria, particularly for the 8-, 16-, and 60-level schemes in combination with different SNR levels. Independent rankings for each criteria will provide important information for design trade-offs between number of levels of quantization and perceptibility factors. The greater the care taken in your rankings by considering differences as a function of each criteria, the more valid and useful the results can be."

EXPERIMENT 2: THRESHOLD AND GAIN

"This is a test of your use of the threshold and gain adjustments available on the Comtal.

"You will view two patterns each at five SNR levels presented in scrambled order. This sequence of ten images will be repeated three times. When an image first appears it will be coded in the normal 6-band, 10-step color code. An amplitude of 1.0 would be white, an amplitude of 0.5 would be yellow.

"Your task is to adjust the threshold and gain with the trackball to improve the color coding over the coding as the image is first presented, if you can. Your objective should be to make significant pattern features of the image as distinctive as possible, relative to the low-level background. This is a subjective judgment which may be difficult to make, particularly on low-SNR images.

"If you press 'R' on the keyboard, you will return to the coding of the image as first presented. You are urged to use this feature freely, particularly on the low-SNR images. Remember, your objective is to bring out the significant pattern features as distinctly as you can.

"A reminder on the use of the trackball to vary threshold and gain: spinning to the left lowers the threshold; to the right raises it. Spinning toward you reduces the gain (spread) of the color scale; spinning away from you increases it.

"When you are satisfied with your threshold and gain adjustments, press the keyboard space bar. This will record your adjustments and advance you to the next image.

"You are urged to take at least five practice trials or more on the image sequence until you sense that your adjustments would be reasonably consistent from one repetition to another."

EXPERIMENT 3: GREY-SCALE VERSUS COLOR AT THE PATTERN THRESHOLD

"This is a test to determine the relative perceptibility of grey-scale versus color coding of acoustic data images. You will view two different patterns on the Comtal. Each of the two will be shown in a test sequence in which the signal strength varies in 11 steps from a maximum of +5.8/2.1 dB (no added noise) to a minimum of -30.0 dB (noise only). The 11 images are labeled 'A' through 'K.' The images will appear on the Comtal four at a time in order of descending signal strength from upper left to lower right. The quadrants are numbered '1' to '4.' When a given target pattern first appears, the image with the highest signal strength in the sequence of 11 images will always appear in the upper left quadrant, so you will be able to note the signal pattern features in the original image. By means of keyboard entries you will be able to 'page' the image sequence, one frame at a time, forward toward the -30.0 dB image or backward toward original image. The 'F' key will move the sequence forward one frame for each press. The 'B' key will reverse the process. The images will shift in position from upper left to upper right and then to lower left and finally lower right. If you press the 'R' key you will return to the original configuration. Pressing the 'N' key will advance you directly to the end of the sequence. To help you keep track of these shifts the letters in parentheses stay with the image.

"Your task in this test will be to judge where in the sequence of 11 images you are at the 'threshold of pattern perceptibility.' You will discover that this is a difficult subjective judgment. Your threshold may vary depending upon the pattern viewed, the type of coding, and the amount of experience you have had. We also expect variability from one subject to another. The more we can reduce the variability from the patterns and subjective factors the more likely we will be able to establish the true difference between the grey-scale and color coding, if any.

"Perhaps the most difficult aspect of the threshold judgment is distinguishing signal peaks from noise peaks when they are approximately equal in strength. To aid you in this judgment, a four-part 'peak locator' cursor has been implemented on the Comtal for this test. Using the trackball, you can position this cursor over any peak desired in any one of the four images and then replicate the cursor by pressing the 'C' on the keyboard. The replicated cursors will automatically be positioned at identical coordinates in the other three images. Use this feature to track either signal peaks or noise peaks as you converge on the image in which the significant pattern features are <u>just</u> perceptible. Several other controls are provided for rapid cursor placement. If you cannot locate the cursor press 'M.' This will position the cursor in the middle of the screen. If you wish to move the cursor rapidly from one quadrant to another, press the space bar.

"Note that the most important aspect of the definition of the threshold is the consistency of your criteria when comparing <u>grey-scale</u> with <u>color</u>. Your consistency in replicating this judgment for the same conditions is also important. The consistency between subjects and between the different patterns is of somewhat lesser importance. To help achieve the desired consistency note that the threshold is defined as that image in which the significant pattern features are still but <u>just barely</u> perceptible. We are arbitrarily defining a threshold which is slightly above the 'true' threshold because the dB difference between the images is fairly great.

"As a further aid to achieving consistency, the test allows you to switch between the grey-scale and color coding for any set of four images. To make this switch, press the 'S' key. You will find this feature particularly valuable for comparing the grey-scale and color after you have decided on the approximate threshold. Note that any cursor placement you make is retained as you switch back and forth. This will aid you in deciding if there is a difference in perceptibility of specific signal features under the two codes. Your judgment of any threshold difference <u>by means of this direct comparison procedure</u> is the <u>essential task</u> in this test. You should think of this task as a <u>unitary</u> comparative judgment. Is there any difference in the perceptibility of the two coding schemes? If so, record the direction and amount of the difference by your responses.

"When you have made your decision, record the thresholds by pressing '1,' '2,' '3,' or '4' on the keyboard when the threshold image is in one of the four quadrants. You must record a threshold for both the grey-scale and the color codes. The order is not important. When you have recorded both thresholds you will advance automatically to the next pattern.

"To ensure that your criteria for the threshold are as consistent as possible, the test is conducted in two phases. In the practice phase you will make your comparative judgments on each of the patterns using exactly the same procedures as in the test itself. However, your responses will not be recorded. Following your practice, the first pattern for the main test will appear automatically. As a result of the judgments made in the practice phase, the

test itself should move rapidly since you are basically confirming and recording your preliminary judgments. When you have made your final threshold determination, you will be sequenced automatically into the second part of the total test which is described below.

"The procedures described above using a 60-level code with six bands and ten intensity steps constitutes the first part of the total test.

"The second part of the test is a repetition of the first in all respects except that now the patterns are coded in 16 levels of grey or color. Again, you must complete the practice phase before your responses are recorded in the test phase.

"The third and final part of the test is a repetition of the first two, now using an 8-level code in grey or color. In summary, the total test consists of threshold determinations on two patterns in both grey-scale and color for three different levels of quantization. This is a total of 2 X 2 X 3 = 12 threshold determinations.

"As a convenience, the use of each of the special keyboard entries and controls is summarized below:

F – Advances the sequence forward toward K, the all noise image.

B – Shifts the sequence backward toward A, the original image.

R – Returns the sequence to the original A, B, C, D configuration.

N – Advances the sequence to the final H, I, J, K configuration.

Trackball – Moves the cursor away from its initial position in the middle of the display.

C – Replicates the cursor in the other three quadrants.

M – Returns the cursor to the middle of the display.

Space Bar – Each press advances the cursor to the same position in the adjacent quadrant.

1, 2, 3, 4, – Uses one of these keys to designate the threshold image.

S – Switches between color and grey-scale."

EXPERIMENT 4: OVERALL RANKING OF GREY-SCALE VERSUS COLOR

"This is a test of the relative perceptibility of grey-scale and color coding measured against the same four perceptibility factors used in the coding scheme study. The experimental design is as follows: 2 (grey/color) X 3 (coding schemes) X 3 (SNR) X 2 (normalized/nonnormalized) X 2 patterns = 72 combinations of variables.

"Your task will be to decide upon rankings upon a scale from 1 (highest) to 7 (lowest) for the color relative to the grey-scale and to record your rankings on the response sheet. These rankings will be made against the four criteria independently. These criteria are

a. Sharpness (perceptibility) of contour edge

b. Overall contrast between adjacent contours

c. Perceptibility of significant pattern features

d. Perceptibility of equivalent amplitudes in widely separated areas.

"If, for a given criteria you cannot perceive a difference between the color and grey codes, give each a rank of 4. If the difference is extreme, give one a rank of 1 and the other a 7. Note that this ranking procedure allows you to express four degrees of difference only: (0,2,4,6), in either direction. To apply such a scale with consistency you will have to develop a subjective standard of differences, i.e., none (4-4), slight (3-5), marked (2-6), and extreme (1-7). Remember, you are to record only one of the four degrees of difference listed above, favoring either the color or grey.

"Since this is a direct comparison of color and grey coding, the test procedure allows you to switch from one to the other at any time by pressing the space bar. You should consider the four criteria, one at a time, switching back and forth between color and grey coding until you have decided and recorded the degree of difference, if any. When you have recorded the rankings against the four criteria, press the 'Return' bar to sequence to the next test condition.

"On the basis of the experiment comparing coding schemes, you should already have in mind the range of differences which can arise when comparing grey and color coding against the four criteria, particularly when the coding scheme and SNR level are varied. It is very important in this experiment that you take advantage of the entire ranking scale available to you, i.e., from 1 to 7. Try to develop a subjective scale in which you assign the ranks of 1-7 to those conditions which you know from previous experience are the most extreme differences you will encounter. Do this even if the perceived difference does not match your conception of 'extreme.' With one end of your scale fixed in mind, assign lesser degrees of difference accordingly, even where the suggested terms 'marked' and 'slight' are not entirely appropriate for you."